

Biotech/ChemLib

62381

From: Gupta, Anish
Sent: Thursday, March 14, 2002 10:34 AM
To: STIC-Biotech/ChemLib

Search Request:

Serial Number 09/544,665
Art Unit: 1653

Date: 3-15-02
Location: 9A13 (mailbox in 9B01)

Please search Seq. ID 1-5.

thank you
anish gupta

Point of Contact:
Mona Smith
Technical Information Specialist
CM1 6A01
Tel: 308-3278

CRFE

Point of Contact:
Mona Smith
Technical Information Specialist
CM1 6A01
Tel: 308-3278

Searcher: M. SMITH
Phone: _____
Location: _____
Date Picked Up: 3/15/02
Date Completed: 3/19/02
Searcher Prep/Review: 15
Clerical: _____
Online time: 15

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: 5
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST(where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

PS Claim 4; Page 37; 64pp; Japanese.
 CC The urinary plasminogen activator (UK) residues 177-184 are used in a
 CC tissue plasminogen activator (t-PA) variant to replace t-PA residues
 CC 294-304. The UK replacement region at least contains the sequence of
 CC AAR74917. t-PA variants, having residues 280-304 substituted with the
 CC corresponding residues from UK, have improved fibrin specificity and
 CC are useful in treatment of thrombosis by reducing adverse reactions of
 CC haemorrhagic tendency.
 XX
 SQ Sequence 8 AA;

Query Match 100.0%; Score 34; DB 16; Length 8;
 Best Local Similarity 100.0%; Pred. No. 4.3e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RHRGS 6
 Db 3 rhrgs 8

RESULT 2
 AAR74919
 ID AAR74919 standard; Peptide: 10 AA.
 XX
 AC AAR74919;
 XX
 DT 13-MAY-1996 (first entry)
 XX
 DE Urea plasminogen activator residues 177-186.
 XX
 KM t-PA; tissue plasminogen activator; urea plasminogen activator; UK;
 KM variant; fibrin specificity; haemorrhagic tendency.
 OS Homo sapiens.
 XX
 PN WO9509908-A1.
 XX
 PD 13-APR-1995.
 XX
 PF 30-SEP-1994; 94WO-JP01629.
 XX
 PR 01-OCT-1993; 93JP-0269983.
 PR 29-JUL-1994; 94JP-0197635.
 XX
 PA (SUMU) SUMITOMO SEIYAKU KK.
 PA (SUMU) SUMITOMO PHARM CO LTD.
 XX
 PI Negoro T, Sato H;
 XX
 DR WPI: 1995-155249/20.
 XX
 PT Tissue plasminogen activator (t-PA) analogue - has increased fibrin
 PT specificity and can be used as an antithrombotic agent.
 XX
 PS Claim 3; Page 37; 64pp; Japanese.
 CC The urinary plasminogen activator (UK) residues 177-186 are used in a
 CC tissue plasminogen activator (t-PA) variant to replace t-PA residues
 CC 294-304. The UK replacement region at least contains the sequence of
 CC AAR74917. t-PA variants, having residues 280-304 substituted with the
 CC corresponding residues from UK, have improved fibrin specificity and
 CC are useful in treatment of thrombosis by reducing adverse reactions of
 CC haemorrhagic tendency.
 XX
 SQ Sequence 10 AA;

Query Match 100.0%; Score 34; DB 16; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.91;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHRGS 6
 Db 3 rhrgs 8
 RESULT 3
 AAR74922
 ID AAR74922 standard; Peptide: 12 AA.
 XX
 AC AAR74922;
 XX
 DT 13-MAY-1996 (first entry)
 XX
 DE Urea plasminogen activator residues 173-183.
 XX
 KM t-PA; tissue plasminogen activator; urea plasminogen activator; UK;
 KM variant; fibrin specificity; haemorrhagic tendency.
 OS Homo sapiens.
 XX
 PN WO9509908-A1.
 XX
 PD 13-APR-1995.
 XX
 PF 30-SEP-1994; 94WO-JP01629.
 XX
 PR 01-OCT-1993; 93JP-0269983.
 PR 29-JUL-1994; 94JP-0197635.
 XX
 PA (SUMU) SUMITOMO SEIYAKU KK.
 PA (SUMU) SUMITOMO PHARM CO LTD.
 XX
 PI Negoro T, Sato H;
 XX
 DR WPI: 1995-155249/20.
 XX
 PT Tissue plasminogen activator (t-PA) analogue - has increased fibrin
 PT specificity and can be used as an antithrombotic agent.
 XX
 PS Claim 6; Page 38; 64pp; Japanese.
 CC The urinary plasminogen activator (UK) residues 173-184 are used in a
 CC tissue plasminogen activator (t-PA) variant to replace t-PA residues
 CC 280-302. The UK replacement region at least contains the sequence of
 CC AAR74917. t-PA variants, having residues 280-304 substituted with the
 CC corresponding residues from UK, have improved fibrin specificity and
 CC are useful in treatment of thrombosis by reducing adverse reactions of
 CC haemorrhagic tendency.
 XX
 SQ Sequence 12 AA;

Query Match 100.0%; Score 34; DB 16; Length 12;
 Best Local Similarity 100.0%; Pred. No. 1.1; 0; Indels 0; Gaps 0;
 Matches 6; Conservative 0; Mismatches 0;
 QY 1 RHRGS 6
 Db 7 rhrgs 12

RESULT 4
 AAR72873
 ID AAR72873 standard; Peptide: 12 AA.
 XX
 AC AAR72873;
 XX
 DT 06-DEC-1995 (first entry)
 XX
 DE Urokinase peptide fragment from amino acids 177-186.
 XX
 KM Urokinase; tissue plasminogen activator; PCR primer; amplification;
 KM thrombosis; blood.

XX OS Synthetic.
 XX PN JP07075580-A.
 XX PD 20-MAR-1995.
 XX PF 28-DEC-1993; 93JP-0354335.
 XX PR 16-JUL-1993; 93JP-0199101.
 XX PA (SUMU) SUMITOMO SEIYAKU KK.
 XX DR WPI: 1995-151486/20.
 XX PT Tissue plasminogen activator (t-PA) mutant comprising sequences
 PT of urinary plasminogen activator - has improved half life in the
 PT blood and may be used to treat thrombosis
 XX Example 2; Page 10; 13pp; Japanese.
 XX PS The peptides AAR72872-3 covering the amino acids (AA) 159-188 or 177-188
 CC of urokinase (UK) were used to replace AA 276-306 of tissue plasminogen
 CC activator (t-PA)(AAR72874) . The modified t-PA also has several N-bound
 CC sugar chain and additional consensus sequences replaced with urokinase
 CC sequences. The replacement was engineered by PCR using the primers
 CC AAG06525-6 to replace the sequences encoding the AA 276-306 in the t-PA
 CC cDNA sequence. The modified t-PA may be used in the treatment of
 CC thrombosis and has an improved half-life in the blood.
 XX SQ Sequence 12 AA;
 Query Match 100.0%; Score 34; DB 16; Length 12;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RHRGGS 6
 DB 3 Thrigs 8
 RESULT 5
 AAR74927
 ID AAR74927 standard; Peptide; 25 AA.
 XX AC AAR74927;
 XX DT 13-MAY-1996 (first entry)
 DE Urea plasminogen activator residues 163-186.
 XX t-PA; tissue plasminogen activator; urea plasminogen activator; UK;
 KM variant; fibrin specificity; haemorrhagic tendency.
 XX OS Homo sapiens.
 XX PN WO9509908-A1.
 XX PD 13-APR-1995.
 XX PF 30-SEP-1994; 94MO-JP01629.
 XX PR 01-OCT-1993; 93JP-0269983.
 XX PR 29-JUL-1994; 94JP-0197635.
 XX PA (SUMU) SUMITOMO SEIYAKU KK.
 XX PA (SUMU) SUMITOMO PHARM CO LTD.
 XX PI Negoro T, Sato H;
 XX DR WPI: 1995-155249/20.

PT Tissue plasminogen activator (t-PA) analogue - has increased fibrin
 PT specificity and can be used as an antithrombotic agent.
 XX PS Claim 1; Page 53; 64pp; Japanese.
 XX CC The urinary plasminogen activator (UK) residues 163-186 are used in a
 CC tissue plasminogen activator (t-PA) variant to replace t-PA residues
 CC 280-304. The UK replacement region at least contains the sequence of
 CC AAR74917. t-PA variants having residues 280-304 substituted with the
 CC corresponding residues from UK, have improved fibrin specificity and
 CC are useful in treatment of thrombosis by reducing adverse reactions of
 CC haemorrhagic tendency.
 XX SQ Sequence 25 AA;
 Query Match 100.0%; Score 34; DB 16; Length 25;
 Best Local Similarity 100.0%; Pred. No. 2.3;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RHRGGS 6
 DB 18 Thrigs 23
 RESULT 6
 AAR42811
 ID AAR42811 standard; peptide; 30 AA.
 XX AC AAR42811;
 XX DT 06-MAY-1994 (first entry)
 DE u-PA (159-188).
 XX Tissue plasminogen activator; t-PA; urinary plasminogen activator;
 KM u-PA; fibrin; fibrinogen; thrombolytic; thrombus; resistance;
 KM plasminogen activation factor inhibiting factor 1.
 XX OS Homo sapiens.
 XX PN WO9320194-A.
 XX PD 14-OCT-1993.
 XX PF 03-APR-1992; 92MO-JP00415.
 XX PR 03-APR-1992; 92MO-JP00415.
 XX PA (SUMU) SUMITOMO PHARM CO LTD.
 XX PI Agui H, Negoro T, Sato H, Sudo Y, Takahashi S;
 XX DR WPI: 1993-336902/42.
 XX PT Tissue plasminogen activator analog improved in fibrin
 PT specificity - obtd. by replacing part of aminoacid sequence with
 PT corresponding aminoacid sequence of urinary plasminogen activator
 XX PS Claim 1-6; Page 25-26; 41pp; Japanese.

CC T-PA analogue has at least part of natural t-PA (276-306)
CC subseq. with corresp. amino acid sequence of u-PA (159-188).
CC The analogue hardly causes activation of plasminogen (and hence
CC decomposition of fibrinogen), while in circulation (blood).
CC It activates plasminogen to cause thrombolysis when it reaches the
CC locus of the thrombus where fibrin is present. It has high fibrin,
CC and therefore thrombus specificity and is not expected to cause
CC general bleeding as side effect. It also is resistant to plasminogen
CC activation factor inhibiting factor 1. Therefore, the analogue will
CC be a safe thrombolytic agent.

SO Sequence 30 AA;

Query Match 100.0%; Score 34; DB 14; Length 30;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRGS 6
| | | | |
DB 21 thrigs 26

RESULT 7

AA93593
ID AA93593 standard; Peptide; 30 AA.

AC AA93593;

DT 13-MAY-1996 (first entry)

DE UK t-PA hybrid peptide CS11 residues 276-306.

KW t-PA; tissue plasminogen activator; urea plasminogen activator; UK;
KM variant; fibrin specificity; haemorrhagic tendency.

OS Synthetic.

PN MO950908-A1.

PD 13-APR-1995.

PF 30-SEP-1994; 94WO-JP01629.

PR 01-OCT-1993; 93JP-026983.

PR 29-JUL-1994; 94JP-0197633.

PA (SUMU) SUMITOMO SEIYAKU KK.
(SUMU) SUMITOMO PHARM CO LTD.

PI Negoro T, Sato H;

DR WPI: 1995-155249/20.

PT Tissue plasminogen activator (t-PA) analogue - has increased fibrin
PT specificity and can be used as an antithrombotic agent.

PS Example 3; Page 50; 64pp; Japanese.

CC Analogues of tissue plasminogen activator (t-PA) have amino acids from
CC position 280-304 replaced by the corresponding sequence of urea
CC plasminogen activator (UK). The UK replacement region at least contains
CC the sequence of AA974917. CS18 comprises a tPA variant (residues
CC 276-306) where residues 294-304 have been replaced by the corresponding
CC UK sequence. The t-PA variants have improved fibrin specificity and are
CC useful in treatment of thrombosis by reducing adverse reactions of
CC haemorrhagic tendency.

SO Sequence 30 AA;

Query Match 100.0%; Score 34; DB 16; Length 30;
Best Local Similarity 100.0%; Pred. No. 2.7;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRGS 6
| | | | |
DB 21 thrigs 26

RESULT 8

AA93594
ID AA93594 standard; Peptide; 30 AA.

AC AA93594;

DT 13-MAY-1996 (first entry)

DE UK t-PA hybrid peptide CS14 residues 276-306.

KW t-PA; tissue plasminogen activator; urea plasminogen activator; UK;
KM variant; fibrin specificity; haemorrhagic tendency.

OS Synthetic.

PN MO950908-A1.

PD 13-APR-1995.

PF 30-SEP-1994; 94WO-JP01629.

PR 01-OCT-1993; 93JP-026983.

PR 29-JUL-1994; 94JP-0197633.

PA (SUMU) SUMITOMO SEIYAKU KK.
(SUMU) SUMITOMO PHARM CO LTD.

PI Negoro T, Sato H;

DR WPI: 1995-155249/20.

PT Tissue plasminogen activator (t-PA) analogue - has increased fibrin
PT specificity and can be used as an antithrombotic agent.

PS Example 4; Page 50; 64pp; Japanese.

CC Analogues of tissue plasminogen activator (t-PA) have amino acids from
CC position 280-304 replaced by the corresponding sequence of urea
CC plasminogen activator (UK). The UK replacement region at least contains
CC the sequence of AA974917. CS18 comprises a tPA variant (residues
CC 276-306) where residues 294-302 have been replaced by the corresponding
CC UK sequence. The t-PA variants have improved fibrin specificity and are
CC useful in treatment of thrombosis by reducing adverse reactions of
CC haemorrhagic tendency.

SO Sequence 30 AA;

Query Match 100.0%; Score 34; DB 16; Length 30;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRGS 6
| | | | |
DB 21 thrigs 26

RESULT 9

AA976947
ID AA976947 standard; Peptide; 30 AA.

AC AA976947;

DT 13-MAY-1996 (first entry)

DE UK t-PA hybrid peptide CS16 residues 276-306.

XX t-PA: tissue plasminogen activator; urea plasminogen activator; UK;
 KW variant; fibrin specificity; haemorrhagic tendency.
 XX Synthetic.
 OS
 XX W09509908-A1.
 PN
 XX 13-APR-1995.
 PD
 XX 30-SEP-1994; 94WO-JP01629.
 PF
 XX 01-OCT-1993; 93JP-0269983.
 PR 29-JUL-1994; 94JP-0197635.
 XX
 PA (SUMU) SUMITOMO SEIYAKU KK.
 PA (SUMU) SUMITOMO PHARM CO LTD.
 XX
 PI Negoro T, Sato H;
 XX
 DR WPI, 1995-155249/20.
 XX
 PT Tissue plasminogen activator (t-PA) analogue - has increased fibrin
 specificity and can be used as an antithrombotic agent.
 XX
 PS Example 6; Page 51; 64pp; Japanese.
 XX
 CC Analogues of tissue plasminogen activator (t-PA) have amino acids from
 CC position 280-304 replaced by the corresponding sequence of urea
 CC plasminogen activator (UK). The UK replacement region at least contains
 CC the sequence of AAR74917. CS15 comprises a tPA variant (residues
 CC 276-306) where residues 290-302 have been replaced by the corresponding
 CC UK sequence. The t-PA variants have improved fibrin specificity and are
 CC useful in treatment of thrombosis by reducing adverse reactions of
 CC haemorrhagic tendency.
 CC
 SO Sequence 30 AA;
 XX
 QY 1 RHRGS 6
 DB 21 thrsgs 26
 XX
 RESULT 10
 ID AAR72872 standard; peptide; 30 AA.
 XX
 AC AAR72872;
 XX
 DT 06-DEC-1995 (first entry)
 XX
 DE Urokinase peptide fragment from amino acids 159-188.
 XX
 KW Urokinase; tissue plasminogen activator; PCR; primer; amplification;
 KW thrombosis; blood.
 XX
 OS Synthetic.
 XX
 PN JP07075580-A.
 PD 20-MAR-1995.
 XX
 PF 28-DEC-1993; 93JP-0354355.
 XX
 PR 16-JUL-1993; 93JP-0199101.
 XX
 PA (SUMU) SUMITOMO SEIYAKU KK.
 XX

DR WPI, 1995-151486/20.
 XX
 XX Tissue plasminogen activator (t-PA) mutant comprising sequences
 PT of urinary plasminogen activator - has improved half life in the
 PT blood and may be used to treat thrombosis
 XX
 PS Claim 1; Page 2; 13pp; Japanese.
 XX
 CC The peptides AAR72872-3 covering the amino acids (AA) 159-188 or 177-188
 CC of urokinase (UK) were used to replace AA 276-306 of tissue plasminogen
 CC activator (t-PA)(AAR72871). The modified t-PA also has several N-bound
 CC sugar chain and additional consensus sequences replaced with urokinase
 CC sequences. The replacement was engineered by PCR using the primers
 CC AA086925-6 to replace the sequences encoding the AA 276-306 in the t-PA
 CC cDNA sequence. The modified t-PA may be used in the treatment of
 CC thrombosis and has an improved half-life in the blood.
 CC
 SO Sequence 30 AA;
 XX
 QY 1 RHRGS 6
 DB 21 thrsgs 26
 XX
 RESULT 11
 ID AAR71666 standard; peptide; 30 AA.
 XX
 AC AAR71666;
 XX
 DT 13-NOV-1995 (first entry)
 XX
 DE Modified urinary plasminogen activator residues 159-188.
 XX
 KW Modified urinary plasminogen activator; residues 159-188;
 KW tissue plasminogen activator; high fibrin specificity;
 KW thrombosis; low systemic bleeding.
 XX
 OS Homo sapiens.
 XX
 PN JP07039374-A.
 PD 10-FEB-1995.
 XX
 PF 28-JUL-1993; 93JP-0207308.
 XX
 PR 28-JUL-1993; 93JP-0207308.
 XX
 PA (SUMU) SUMITOMO SEIYAKU KK.
 XX
 DR WPI, 1995-117856/16.
 XX
 PT Novel modified tissue plasminogen activator - contains a urinary
 PT plasminogen activator peptide fragment
 XX
 PS Claim 1; Page 8; 10pp; Japanese.
 XX
 CC AAR79104 is the wild type plasminogen activator (t-PA), residues
 CC 276-306. By replacing these residues with the urinary
 CC plasminogen activator (UK) sequences described in AAR71662-R71666,
 CC the fibrinogen specificity of t-PA is increased. The modified
 CC t-PA can be used in the treatment of thrombosis and low systemic
 CC bleeding.
 CC
 SO Sequence 30 AA;
 XX
 QY Query Match 100.0%; Score 34; DB 16; Length 30;
 Best Local Similarity 100.0%; Pred. No. 2.7; Indels 0; Gaps 0;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRGS 6
|||
DB 21 thrigs 26

RESULT 12

AAR24949
ID AAR24949 standard; peptide: 32 AA.

AC AAR24949;

DT 15-DEC-1992 (first entry)

DE Urokinase fragment.

KW Tissue plasminogen activator; analogue; mutant; PAI; inhibitor; thrombosis.

OS Synthetic.

PN JP04144682-A.

PD 19-MAY-1992.

PF 05-OCT-1990; 90JP-0268816.

PR 05-OCT-1990; 90JP-0268816.

PA (SUMU) SUMITOMO SEIYAKU KK.

DR WPI: 1992-214121/26.

DR N-PSDB; AAQ25737.

PT New tissue plasminogen activator analogue - is resistant to PA inhibitor but retains properties of natural t-PA

PS Claim 1; Fig 6; 13pp; Japanese.

CC The urokinase sequence was used in a construct for a novel tissue plasminogen activator (tPA) having amino acids 276-306 of tPA substituted by this urokinase sequence. The tPA analogue has resistance to inhibition by plasminogen activator inhibitor (PAI) while retaining its tPA activity. The analogue can be used in a drug for the treatment of thrombosis.

SQ Sequence 32 AA;

Query Match 100.0%; Score 34; DB 13; Length 32;

Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRGS 6
|||
DB 21 thrigs 26

RESULT 13

AAR64211
ID AAR64211 standard; peptide: 39 AA.

AC AAR64211;

DT 13-MAY-1996 (first entry)

DE Urinary plasminogen activator residues 150-188.

KW t-PA; tissue plasminogen activator; urinary plasminogen activator; UK; variant; fibrin specificity; haemorrhagic tendency.

XX

OS Homo sapiens.

PN JP06327473-A.

PD 29-NOV-1994.

PF 19-MAY-1993; 93JP-0141351.

PR 19-MAY-1993; 93JP-0141351.

PA (SUMU) SUMITOMO SEIYAKU KK.

DR WPI: 1995-047896/07.

PT Novel t-PA variant containing urinary plasminogen activator residues - has improved fibrin specificity to reduce adverse reaction of haemorrhagic tendency

PS Claim 3; Page 8-9; 15pp; Japanese.

CC The urinary plasminogen activator (UK) residues 150-188 are used in a tissue plasminogen activator (t-PA) variant to replace t-PA residues 266-306. t-PA variants, having residues 276-306 substituted with the corresponding residues from UK, have improved fibrin specificity. They are useful in reducing adverse reactions of haemorrhagic tendency.

SQ Residues 266-275 and/or 307-321 are opt. substd.

Sequence 39 AA;

Query Match 100.0%; Score 34; DB 16; Length 39;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRGS 6
|||
DB 30 thrigs 35

RESULT 14

AAR64212
ID AAR64212 standard; peptide: 45 AA.

AC AAR64212;

DT 13-MAY-1996 (first entry)

DE Urinary plasminogen activator residues 159-203.

KW t-PA; tissue plasminogen activator; urinary plasminogen activator; UK; variant; fibrin specificity; haemorrhagic tendency.

OS Homo sapiens.

PN JP06327473-A.

PD 29-NOV-1994.

PF 19-MAY-1993; 93JP-0141351.

PR 19-MAY-1993; 93JP-0141351.

PA (SUMU) SUMITOMO SEIYAKU KK.

DR WPI: 1995-047896/07.

PT Novel t-PA variant containing urinary plasminogen activator residues - has improved fibrin specificity to reduce adverse reaction of haemorrhagic tendency

PS Claim 4; Page 9; 15pp; Japanese.

XX The urinary plasminogen activator (UK) residues 159-203 are used in a

CC tissue plasminogen activator (t-PA) variant to replace t-PA residues
 CC 266-306. t-PA variants, having residues 276-321 substituted with the
 CC corresponding residues from UK, have improved fibrin specificity. They
 CC are useful in reducing adverse reactions of haemorrhagic tendency.
 CC Residues 266-275 and/or 307-321 are opt. substd..

XX Sequence 45 AA;

Query Match 100.0%; Score 34; DB 16; Length 45;
 Best Local Similarity 100.0%; Pred. No. 4.1;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RHRCGS 6
 |||||
 DB 21 rhrgs 26

RESULT 15

AA064213
 ID AA064213 standard; Peptide: 54 AA.

XX AA064213;

AC AA064213;

DT 13-MAY-1996 (first entry)

DE Urinary plasminogen activator residues 150-203.

KW t-PA; tissue plasminogen activator; urinary plasminogen activator; UK;
 variant; fibrin specificity; haemorrhagic tendency.

OS Homo sapiens.

PN JP06327473-A.

PD 29-NOV-1994.

PF 19-MAY-1993; 93JP-0141351.

PR 19-MAY-1993; 93JP-0141351.

PA (SUMU) SUMITOMO SEIYAKU KK.

DR WPI; 1995-047896/07.

PT Novel t-PA variant containing urinary plasminogen activator residues
 PT - has improved fibrin specificity to reduce adverse reaction of
 PT haemorrhagic tendency

PS Claim 5; Page 9; 15pp; Japanese.

CC The urinary plasminogen activator (UK) residues 150-203 are used in a
 CC tissue plasminogen activator (t-PA) variant to replace t-PA residues
 CC 266-321. t-PA variants, having residues 276-306 substituted with the
 CC corresponding residues from UK, have improved fibrin specificity. They
 CC are useful in reducing adverse reactions of haemorrhagic tendency.
 CC Residues 266-275 and/or 307-321 are opt. substd..

XX Sequence 54 AA;

Query Match 100.0%; Score 34; DB 16; Length 54;
 Best Local Similarity 100.0%; Pred. No. 4.9;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHRCGS 6
 |||||
 DB 30 rhrgs 35

Search completed: March 18, 2002, 09:48:55
 Job time: 364 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OW protein - protein search, using sw model

Run on: March 18, 2002, 09:45:56 ; Search time 37.77 Seconds
(without alignments)
3.575 Million cell updates/sec

Title: US-09-544-665-1
Sequence: 1 RHRCGS 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503297 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCRTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/Dackfiles.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	100.0	6	US-07-942-157A-1	Sequence 1, Appl1
2	34	100.0	253	US-08-944-483-73	Sequence 73, Appl1
3	34	100.0	254	US-08-560-098A-49	Sequence 49, Appl1
4	34	100.0	306	US-08-560-098A-45	Sequence 45, Appl1
5	34	100.0	331	US-08-560-098A-46	Sequence 46, Appl1
6	34	100.0	365	US-08-093-741-83	Sequence 83, Appl1
7	34	100.0	365	US-08-720-012-83	Sequence 83, Appl1
8	34	100.0	393	US-08-560-098A-44	Sequence 44, Appl1
9	34	100.0	393	US-08-967-024C-24	Sequence 24, Appl1
10	34	100.0	393	US-08-967-024C-25	Sequence 25, Appl1
11	34	100.0	411	US-08-087-163-1	Sequence 1, Appl1
12	34	100.0	411	US-08-286-748B-18	Sequence 18, Appl1
13	34	100.0	411	US-08-153-799-18	Sequence 48, Appl1
14	34	100.0	411	US-08-560-098A-48	Sequence 48, Appl1
15	34	100.0	411	US-09-181-815-1	Sequence 1, Appl1
16	34	100.0	430	US-07-942-157A-3	Sequence 3, Appl1
17	34	100.0	430	519569-2	Sequence 3, Appl1
18	34	100.0	431	5188829-1	Patent No. 5188829
19	34	100.0	432	US-08-560-098A-47	Sequence 47, Appl1
20	31	91.2	23	US-08-667-809B-4	Sequence 4, Appl1
21	30	88.2	434	US-08-337-602-3	Sequence 3, Appl1
22	30	88.2	434	US-08-558-135-3	Sequence 3, Appl1
23	30	88.2	455	US-08-272-255-14	Sequence 14, Appl1
24	30	88.2	455	PCT-US95-08565-14	Sequence 14, Appl1
25	30	88.2	566	US-07-955-905A-2	Sequence 2, Appl1
26	30	88.2	566	US-07-955-905A-22	Sequence 22, Appl1
27	30	88.2	1096	US-09-346-237-6	Sequence 6, Appl1

28	30	88.2	1323	1	US-08-026-138E-4	Sequence 4, Appl1
29	30	88.2	1336	2	US-08-231-193A-58	Sequence 58, Appl1
30	30	88.2	1336	2	US-08-486-273A-58	Sequence 58, Appl1
31	30	88.2	1336	3	US-08-940-086A-58	Sequence 58, Appl1
32	30	88.2	2237	1	US-08-455-543A-48	Sequence 48, Appl1
33	30	88.2	2237	2	US-08-223-305C-48	Sequence 2, Appl1
34	30	88.2	2337	3	US-08-713-118-2	Sequence 2, Appl1
35	30	88.2	2337	4	US-09-453-007-2	Sequence 2, Appl1
36	30	88.2	2339	1	US-08-455-543A-47	Sequence 47, Appl1
37	30	88.2	2339	2	US-08-223-305C-47	Sequence 47, Appl1
38	28	82.4	50	4	US-08-817-787-22	Sequence 22, Appl1
39	28	82.4	124	1	US-08-441-629-5	Sequence 5, Appl1
40	28	82.4	124	3	US-08-776-207-5	Sequence 5, Appl1
41	28	82.4	124	5	PCT-US95-09172-5	Sequence 5, Appl1
42	28	82.4	125	2	US-08-891-848-16	Sequence 16, Appl1
43	28	82.4	125	3	US-08-875-811-11	Sequence 11, Appl1
44	28	82.4	125	6	5171845-2	Patent No. 5171845
45	28	82.4	226	3	US-09-195-286-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-07-942-157A-1
Sequence 1, Application US/07942157A
Patent No. 5648253
GENERAL INFORMATION:
APPLICANT: Wei, Cha-Mer
TITLE OR INVENTION: Inhibitor-Resistant Urokinase
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 1100 Peachtree Street Suite 2800
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-4530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/942,157A
FILING DATE: 19920908
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/631673
FILING DATE: 20-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Padst, Patricia L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: TS1108Cont.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)815-6508
TELEFAX: (404)815-6555
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-07-942-157A-1
Query Match 100.0%; Score 34; DB 1; Length 6;

Best Local Similarity 100.0%; Pred. NO. 1.6e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHRGS 6
111111
DB 1 RHRGS 6

RESULT 2
US-08-944-483-73

; Sequence 73, Application US/08944483
; Patent No. 6232456

; GENERAL INFORMATION:

; APPLICANT: COHEN, MAURICE

; APPLICANT: COLPITTS, TRACEY L.

; APPLICANT: FRIEDMAN, PAULA N.

; APPLICANT: GRANADOS, EDWARD N.

; APPLICANT: KLAS, MICHAEL R.

; APPLICANT: RUSSELL, JOHN C.

; APPLICANT: STEWART, KENT D.

; APPLICANT: STROUPE, STEVEN D.

; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS

; TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES

; NUMBER OF SEQUENCES: 76

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Abbott Laboratories

; STREET: 100 Abbott Park Road

; CITY: Abbott Park

; STATE: IL

; COUNTRY: USA

; ZIP: 60064-3500

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/944,483

; FILING DATE:

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Becker, Cheryl L.

; REGISTRATION NUMBER: 35,441

; REFERENCE/DOCKET NUMBER: 6183.US.01

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 847/935-1729

; TELEFAX: 847/938-2623

; TELEX:

; INFORMATION FOR SEQ ID NO: 73:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 253 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: No. 6232456e

; US-08-944-483-73

Query Match 100.0%; Score 34; DB 4; Length 253;
Best Local Similarity 100.0%; Pred. No. 17;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHRGS 6
111111
DB 21 RHRGS 26

RESULT 3
US-08-560-098A-49

; Sequence 49, Application US/08560098A
; Patent No. 5976841

; GENERAL INFORMATION:

; APPLICANT: WENDEL, Stephan

; APPLICANT: HEINZEL-WIELAND, Regina

; APPLICANT: STEFFENS, Gerd Josef

; TITLE OF INVENTION: Proteins having fibrinolytic and

; TITLE OF INVENTION: Coagulation-inhibiting Properties

; NUMBER OF SEQUENCES: 60

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Evenson, McKown, Edwards & Lenahan

; STREET: 1200 G Street, N.W., Suite 700

; CITY: Washington

; STATE: DC

; COUNTRY: USA

; ZIP: 20005

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/560,098A

; FILING DATE: 17-NOV-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: P 44 40 892.7

; FILING DATE: 17-NOV-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: EVANS, Joseph D.

; REGISTRATION NUMBER: 26,269

; REFERENCE/DOCKET NUMBER: 148/42448

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 628-8800

; TELEFAX: (202) 628-8844

; INFORMATION FOR SEQ ID NO: 49:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 254 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-560-098A-49

Query Match 100.0%; Score 34; DB 2; Length 254;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHRGS 6
111111
DB 22 RHRGS 27

RESULT 4
US-08-560-098A-45

; Sequence 45, Application US/08560098A
; Patent No. 5976841

; GENERAL INFORMATION:

; APPLICANT: WENDEL, Stephan

; APPLICANT: HEINZEL-WIELAND, Regina

; APPLICANT: STEFFENS, Gerd Josef

; TITLE OF INVENTION: Proteins having fibrinolytic and

; TITLE OF INVENTION: Coagulation-inhibiting Properties

; NUMBER OF SEQUENCES: 60

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Evenson, McKown, Edwards & Lenahan

; STREET: 1200 G Street, N.W., Suite 700

; CITY: Washington

; STATE: DC

; COUNTRY: USA

; ZIP: 20005

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,098A
FILING DATE: 17-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 40 892.7
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42448
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8844
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 306 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-560-098A-45

Query Match 100.0%; Score 34; DB 2; Length 306;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHRGGS 6
Db 74 RHRGGS 79

RESULT 5
US-08-560-098A-46
Sequence 46, Application US/08560098A
Patent No. 5976841
GENERAL INFORMATION:
APPLICANT: WNEEDT, Stephan
APPLICANT: HEINZEL-WIELAND, Regina
APPLICANT: STEFFENS, Gerd Josef
TITLE OF INVENTION: Proteins having fibrinolytic and
TITLE OF INVENTION: Coagulation-inhibiting Properties
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,098A
FILING DATE: 17-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 40 892.7
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42448
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8844
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:

LENGTH: 331 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-560-098A-46

Query Match 100.0%; Score 34; DB 2; Length 331;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHRGGS 6
Db 74 RHRGGS 79

RESULT 6
US-08-093-741-83
Sequence 83, Application US/08093741
Patent No. 5681721
GENERAL INFORMATION:
APPLICANT: STEFFENS, GERD J.
APPLICANT: WNEEDT, STEPHAN
APPLICANT: SCHNEIDER, JOHANNES
APPLICANT: HEINZEL-WIELAND, REGINA
APPLICANT: SAUNDERS, DEREK J.
TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
TITLE OF INVENTION: INHIBITING EFFECT
NUMBER OF SEQUENCES: 83
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington, D.C.
COUNTRY: U.S.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/093,741
FILING DATE: 20-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P43 23 754.1
FILING DATE: 15-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/41345
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)628-8844
TELEFAX: (202)628-8844
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-093-741-83

Query Match 100.0%; Score 34; DB 1; Length 365;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHRGGS 6
Db 133 RHRGGS 138

RESULT 7
US-08-720-012-83

Sequence 83, Application US/08720012
Patent No. 5747291

GENERAL INFORMATION:

APPLICANT: STEFFENS, GERD J.

APPLICANT: WENNDT, STEPHAN

APPLICANT: SCHNEIDER, JOHANNES

APPLICANT: HEINZEL-WIELAND, REGINA

TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN

TITLE OF INVENTION: INHIBITING EFFECT

NUMBER OF SEQUENCES: 83

CORRESPONDENCE ADDRESS:

ADDRESSEE: Evenson, McKeown, Edwards & Lenahan

STREET: 1200 G Street, N. W., Suite 700

CITY: Washington, D.C.

COUNTRY: U.S.

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/720,012

FILING DATE: 27-SEP-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/093,741

FILING DATE: 20-JUL-1993

APPLICATION NUMBER: DE P43 23 754.1

FILING DATE: 15-JUL-1993

ATTORNEY/AGENT INFORMATION:

NAME: EVANS, Joseph D.

REGISTRATION NUMBER: 26,269

REFERENCE/DOCKET NUMBER: 148/41345

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)628-8800

TELEFAX: (202)628-8844

INFORMATION FOR SEQ ID NO: 83:

SEQUENCE CHARACTERISTICS:

LENGTH: 365 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-08-720-012-83

Query Match

Best Local Similarity 100.0%; Score 34; DB 1; Length 365;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRRGS 6

DB 133 RHRRGS 138

RESULT 8

US-08-560-098A-44

Sequence 44, Application US/08560098A

Patent No. 5976841

GENERAL INFORMATION:

APPLICANT: WENNDT, Stephan

APPLICANT: HEINZEL-WIELAND, Regina

APPLICANT: STEFFENS, Gerd Josef

TITLE OF INVENTION: proteins having fibrinolytic and

TITLE OF INVENTION: Coagulation-inhibiting Properties

NUMBER OF SEQUENCES: 60

CORRESPONDENCE ADDRESS:

ADDRESSEE: Evenson, McKeown, Edwards & Lenahan

STREET: 1200 G Street, N.W., Suite 700

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/560,098A

FILING DATE: 17-NOV-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: P 44 40 892.7

FILING DATE: 17-NOV-1994

ATTORNEY/AGENT INFORMATION:

NAME: EVANS, Joseph D.

REGISTRATION NUMBER: 26,269

REFERENCE/DOCKET NUMBER: 148/42448

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 628-8800

TELEFAX: (202) 628-8844

INFORMATION FOR SEQ ID NO: 44:

SEQUENCE CHARACTERISTICS:

LENGTH: 393 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-560-098A-44

Query Match

Best Local Similarity 100.0%; Score 34; DB 2; Length 393;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRRGS 6

DB 134 RHRRGS 139

RESULT 9

US-08-967-024C-24

Sequence 24, Application US/08967024C

Patent No. 6133011

GENERAL INFORMATION:

APPLICANT: WENNDT, Stephan

APPLICANT: STEFFENS, Gerd Josef

APPLICANT: JANCOCHA, Elke

APPLICANT: HEINZEL-WIELAND, Regina

TITLE OF INVENTION: Chimeric Proteins having fibrinolytic

TITLE OF INVENTION: Chimeric Proteins having fibrinolytic

NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESS:

ADDRESSEE: Evenson, McKeown, Edwards & Lenahan

STREET: 1200 G Street, N.W., Suite 700

CITY: Washington

STATE: DC

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/967,024C

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: P 44 42 665.8

FILING DATE: 30-NOV-1994

ATTORNEY/AGENT INFORMATION:

NAME: EVANS, Joseph D.

REGISTRATION NUMBER: 26,269

REFERENCE/DOCKET NUMBER: 148/42444

TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-967-024C-24

Query Match 100.0%; Score 34; DB 4; Length 393;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRGS 6
DB 134 RHRGS 139

RESULT 10
US-08-967-024C-25
Sequence 25, Application US/08967024C
Patent No. 6133011
GENERAL INFORMATION:
APPLICANT: WENNDT, Stephan
APPLICANT: STEPPENS, Gerd Josef
APPLICANT: JANCOCH, Elke
APPLICANT: HEINZEL-WIELAND, Regina
TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,024C
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 42 665.8
FILING DATE: 30-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 146/42444
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-967-024C-25

Query Match 100.0%; Score 34; DB 4; Length 393;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 RHRGS 6

DB 134 RHRGS 139

RESULT 11
US-08-087-163-1
Sequence 1, Application US/08087163
Patent No. 5472692
GENERAL INFORMATION:
APPLICANT: Liu, Jian-Ning
APPLICANT: Gurevich, Victor
TITLE OF INVENTION: PRO-UROKINASE MUTANTS
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" diskette, 1.44 MB
COMPUTER: IBM PS/2 Model 502 or 555X
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: Wordperfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/087,163
FILING DATE: 07/02/93
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, J. Peter
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 04353/003001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 411
TYPE: amino acid
STRANDEDNESS: N/A
TOPOLOGY: N/A
US-08-087-163-1

Query Match 100.0%; Score 34; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRGS 6
DB 179 RHRGS 184

RESULT 12
US-08-286-748B-18
Sequence 18, Application US/08286748B
Patent No. 5759542
GENERAL INFORMATION:
APPLICANT: Victor Gurevich
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY
OF DRUGS BY PLATELETS FOR THE TREATMENT OF
TITLE OF INVENTION: CARDIOVASCULAR AND OTHER DISEASES
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts

COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 555X
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/786,748B
FILING DATE: August 5, 1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: J. Peter Fassee
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 04547/013001
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 411
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-286-748B-18

Query Match 100.0%; Score 34; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHRGGS 6
Db 179 RHRGGS 184

RESULT 13
US-08-153-799-18
Sequence 18, Application US/08153799
Patent No. 576683
GENERAL INFORMATION:
APPLICANT: Ballance, David J
APPLICANT: Goodey, Andrew R
TITLE OF INVENTION: Polypeptides
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: R Hain Swope, BOC Health Care Inc
STREET: 100 Mountain Avenue
CITY: Murray Hill
STATE: New Jersey
COUNTRY: USA
ZIP: 07974
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/153,799
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847975
FILING DATE: 06-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 8909916.2
FILING DATE: 29-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB90/00650

FILING DATE: 26-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/775952
FILING DATE: 29-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Swope, R Hain
REGISTRATION NUMBER: 24864
REFERENCE/DOCKET NUMBER: 928832
TELEPHONE: (908) 665 2400
TELEFAX: (908) 771 6159
TELEX: 219484
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 411 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-153-799-18

Query Match 100.0%; Score 34; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHRGGS 6
Db 179 RHRGGS 184

RESULT 14
US-08-560-098A-48
Sequence 48, Application US/08560098A
Patent No. 5976841
GENERAL INFORMATION:
APPLICANT: WENNDT, Stephan
APPLICANT: HEINZEL-WIELAND, Regina
TITLE OF INVENTION: Proteins having fibrinolytic and
TITLE OF INVENTION: Coagulation-inhibiting Properties
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKee, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,098A
FILING DATE: 17-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 40 892.7
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42448
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 411 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-560-098A-48

Query Match 100.0%; Score 34; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHRGS 6
|||||
DB 179 RHRGS 184

RESULT 15
US-09-181-816-1
; Sequence 1, Application US/09181816
; Patent No. 6277818
; GENERAL INFORMATION:
; APPLICANT: MAZAR, Andrew P.
; TITLE OF INVENTION: CYCLIC PEPTIDE LIGANDS THAT TARGET UROKINASE
; FILE REFERENCE: 329042000300 SIDN 1-7
; CURRENT APPLICATION NUMBER: US/09/181,816
; NUMBER OF SEQ. ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ. ID NO. 1
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-181-816-1

Query Match 100.0%; Score 34; DB 4; Length 411;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHRGS 6
|||||
DB 179 RHRGS 184

Search completed: March 18, 2002, 09:49:40
Job time: 224 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:46:31 ; Search time 41.98 Seconds

(Without alignments)
10.887 Million cell updates/sec

Title: US-09-544-665-1

Perfect score: 34

Sequence: 1 RHRGS 6

Scoring table: BLOSUM62

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	100.0	431	1 UKHU	u-plasminogen acti
2	34	100.0	433	1 UKBAY	u-plasminogen acti
3	34	100.0	433	1 JN0560	u-plasminogen acti
4	34	100.0	442	1 UKPG	u-plasminogen acti
5	34	100.0	597	2 T09681	adenosylhomocystei
6	34	100.0	1217	2 S52714	sericthib silkmo
7	34	100.0	1332	2 T23024	hypothetical prote
8	31	91.2	64	2 S32789	toxin II-14 - scor
9	31	91.2	86	2 JN0671	Na ⁺ -channel-blocki
10	31	91.2	156	2 C84688	probable C2H2-type
11	31	91.2	232	2 E83606	hypothetical prote
12	31	91.2	310	2 T43147	hypothetical prote
13	31	91.2	376	2 J70800	probable ABC trans
14	31	91.2	431	2 J70800	carbazole dioxygen
15	31	91.2	501	2 T14339	sucrose-proton tra
16	31	91.2	508	2 E64089	probable transport
17	30	88.2	53	2 S10355	T-cell receptor be
18	30	88.2	88	2 F85717	hypothetical prote
19	30	88.2	88	2 A85743	hypothetical prote
20	30	88.2	111	2 T49404	hypothetical prote
21	30	88.2	118	2 A72316	conserved hypotet
22	30	88.2	130	2 A72650	hypothetical prote
23	30	88.2	163	2 T36985	hypothetical prote
24	30	88.2	179	2 S63610	probable prophage
25	30	88.2	266	1 R5NR2D	ribosomal protein
26	30	88.2	273	1 R5R2	ribosomal protein
27	30	88.2	273	1 R5ZM2	ribosomal protein
28	30	88.2	273	1 B71355	probable ribosomal
29	30	88.2	274	1 R5NR2	ribosomal protein

30	30	88.2	274	2 S78397	ribosomal protein
31	30	88.2	275	2 T11810	ribosomal protein
32	30	88.2	275	2 T07361	ribosomal protein
33	30	88.2	275	2 S78273	ribosomal protein
34	30	88.2	276	2 S77499	ribosomal protein
35	30	88.2	276	2 B54547	ribosomal protein
36	30	88.2	276	2 A72250	ribosomal protein
37	30	88.2	276	2 T07531	ribosomal protein
38	30	88.2	284	2 F72055	ribosomal protein
39	30	88.2	284	2 C71597	probable l2 riboso
40	30	88.2	284	2 A86571	l2 ribosomal prote
41	30	88.2	284	2 A81661	ribosomal protein
42	30	88.2	284	2 C83237	probable transcript
43	30	88.2	286	1 R5SP2	ribosomal protein
44	30	88.2	304	2 G70300	ribosomal protein
45	30	88.2	342	1 JN0542	head protein gp7 -

ALIGNMENTS

RESULT 1

UKHU
u-plasminogen activator (EC 3.4.21.73) precursor [validated] - human
N:Alternate names: cellular plasminogen activator; urokinase; urokinase-type plasmin
N:Contains: urokinase-type plasminogen activator chain A; urokinase-type plasminogen
in form
C:Species: Homo sapiens (man)
C>Date: 17-Dec-1982 #sequence revision 04-Dec-1986 #text change 15-Sep-2000
C/Accession: A00931; I52209; J70102; A37561; I38102; S65783; A37562; A37563; A37564;
R:Riccio, A.; Grimaldi, G.; Verle, P.; Sebastio, G.; Boast, S.; Blasl, F.
Nucleic Acids Res. 13, 2759-2771, 1985
A>Title: The human urokinase-plasminogen activator gene and its promoter.
A:Reference number: A00931; MUID:85215647
A/Accession: A00931
A:Molecule type: DNA
A:Residues: 1-431 <RC>
A/Cross-references: GB:X02419; NID:937601; PIDN:CA26268.1; PID:91834524
A/Note: the authors translated the codon ATG for residue 214 as Ile
R:Nagamine, T.; Pearson, D.; Gratian, M.
Biochem. Biophys. Res. Commun. 132, 563-569, 1985
A>Title: Exon-intron boundary sliding in the generation of two mRNAs coding for porc
A:Reference number: I52209; MUID:86050639
A/Accession: I52209
A/Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 145-161 <NAG1>
A/Cross-references: GB:K03027; NID:9340174; PIDN:AAA61257.1; PID:9340175
R:Nagel, M.; Hiratsugu, R.; Kanada, T.; Hayasuke, N.; Arimura, H.; Nishida, M.; Suyu
Gene 36, 183-188, 1985
A>Title: Molecular cloning of cDNA coding for human preprourokinase.
A:Reference number: J70102; MUID:86050634
A/Accession: J70102
A:Molecule type: mRNA
A:Residues: 1-213, 'I', 215-431 <NAG2>
A/Cross-references: GB:K03226; NID:9340155; PIDN:AC97138.1; PID:9340158; GB:DO0244;
R:Verle, P.; Stoppelli, M.P.; Galeffi, P.; Di Nocera, P.; Blasl, F.
Proc. Natl. Acad. Sci. U.S.A. 81, 4727-4731, 1984
A>Title: Identification and primary sequence of an unspliced human urokinase poly(A)
A:Reference number: A37561; MUID:84272706
A/Accession: A37561
A:Molecule type: mRNA
A:Residues: 66-431 <VER>
A/Cross-references: GB:000244; NID:9220138
R:Jacobs, P.; Cravador, A.; Lortiau, R.; Brockly, F.; Colau, B.; Chuchane, P.; van El
DNA 4, 139-146, 1985
A>Title: Molecular cloning, sequencing, and expression in *Escherichia coli* of human
A:Reference number: I38102; MUID:85203359
A/Accession: I38102
A/Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-150, 'W', 152-213, 'I', 215-385, 'C', 387-429, 'V', 431 <JAC>
A/Cross-references: EMBL:X02760; NID:935297; PIDN:CAA26535.1; PID:935298

R.Yoshimoto, M.; Ushiyama, Y.; Sakai, M.; Tamaki, S.; Hara, H.; Takahashi, K.; Sawasaki, Biochim. Biophys. Acta 1293, 83-89, 1996
 A:Title: Characterization of single chain urokinase-type plasminogen activator with a no
 A:Reference number: S65783; MUID:96186279
 A:Accession: S65783
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 21-140 'L', 142-213 'I', 215-431 <YOS>
 A:Cross-references: EMBL:D1143; NID:91311467; PIDN:BA01919.1; PID:q1199928
 R:Gunter, M.A.; Steffens, G.J.; Oetting, F.; Kim, S.M.A.; Franks, E.; Flohe, L.
 Hoppe-Seyler's Z. Physiol. Chem. 363, 1155-1165, 1982
 A:Title: The primary structure of high molecular mass urokinase from human urine.
 A:Reference number: A37562; MUID:83055084
 A:Accession: A37562
 A:Molecule type: protein
 A:Residues: 21-177 <GUN>
 R:Schaller, J.; Nick, R.; Rickli, E.E.; Gillesen, D.; Lergier, W.; Studer, R.O.
 Eur. J. Biochem. 125, 251-257, 1982
 A:Title: Human low-molecular-weight urinary urokinase. Partial characterization and prel
 A:Reference number: A37563; MUID:83003608
 A:Accession: A37563
 A:Molecule type: protein
 A:Residues: 155-176; 179-193 'T', 195 'T', 197-224 <SCH>
 R:Steffens, G.J.; Gunzler, W.A.; Oetting, F.; Franks, E.; Flohe, L.
 Hoppe-Seyler's Z. Physiol. Chem. 363, 1043-1058, 1982
 A:Title: The complete amino acid sequence of low molecular mass urokinase from human ur
 A:Reference number: A37564; MUID:83055099
 A:Accession: A37564
 A:Molecule type: protein
 A:Residues: 158-410 <STE>
 R:Kretzer, E.J.; Buho, A.; Menon, G.; Sarin, V.K.
 Biochem. Biophys. Res. Commun. 171, 401-406, 1990
 A:Title: Carbohydrate composition and presence of a fucose-protein linkage in recombinan
 A:Reference number: A35689; MUID:90365737
 A:Accession: A35689
 A:Molecule type: protein
 A:Residues: 21-30, 'X', 32, 'X', 34-38, 'X', 40-43 <KEN>
 A:Note: Identification of a fucose and attempt to determine its attachment site
 R:Rabdan, S.A.; Desjardins, J.; Bell, A.W.; Banyville, D.; Mazur, A.; Henkin, J.; Goltz
 Biochem. Biophys. Res. Commun. 173, 1058-1064, 1990
 A:Title: An amino-terminal fragment of urokinase isolated from a prostate cancer cell li
 A:Reference number: A36697; MUID:91095529
 A:Accession: A36697
 A:Molecule type: protein
 A:Residues: 21-34 <RAB>
 R:Li, X.; Bokman, A.M.; Llinas, M.; Smith, R.A.G.; Dobson, C.M.
 Submitted to the Brookhaven Protein Data Bank, July 1993
 A:Reference number: A51255; PDB:1XDU
 A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
 R:Li, X.; Smith, R.A.G.; Dobson, C.M.
 Biochemistry 31, 9562-9571, 1992
 A:Title: Sequential (1)H NMR assignments and secondary structure of the kringle domain f
 A:Reference number: A44375; MUID:93003110
 A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR
 R: Hansen, A.P.; Petros, A.M.; Keadow, R.P.; Nettesheim, D.G.; Mazur, A.P.; Olejniczak,
 submitted to the Brookhaven Protein Data Bank, January 1994
 A:Reference number: A66822; PDB:1URK
 A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
 R:Spraggon, G.S.; Phillips, C.; Nowak, U.K.; Ponting, C.P.; Saunders, D.; Dobson, C.M.;
 submitted to the Brookhaven Protein Data Bank, July 1995
 A:Reference number: A66058; PDB:1LWM
 A:Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 168-175; 179-426
 C:Comment: This enzyme is found in urine in a high molecular mass form, consisting of A
 C:Comment: Urokinase-type plasminogen activator proteolytically activates plasminogen,
 C:Genetics:
 A:Gene: GDB:PLNU
 A:Cross-references: GDB:119497; OMIM:191840
 A:Map position: 10q24-10q24
 A:Intons: 193; 29/1; 65/1; 123/2; 154/1; 227/2; 277/1; 324/1; 373/3
 C:Function:
 A:Description: proteolytically activates plasminogen
 A:Pathway: fibrinolysis
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try

C:Keywords: fibrinolysis; glycoprotein; heterodimer; hydrolase; kringle; serine prote
 F:1-10/Domain: signal sequence #status predicted <Sig>
 F:21-431/Product: urokinase-type plasminogen activator, single chain form #status pre
 F:21-177/Product: urokinase-type plasminogen activator chain A #status experimental <
 F:31-62/Domain: EGF homology <EGF>
 F:70-151/Domain: kringle homology <KR>
 F:156-177/Product: urokinase-type plasminogen activator chain A1 #status experimental
 F:179-431/Product: urokinase-type plasminogen activator chain B #status experimental
 F:179-439/Domain: trypsin homology <TRY>
 F:31-39; 33-51; 53-62; 70-151; 91-153; 122-146; 168-299; 209-225; 217-288; 313-382; 345-361; 372
 F:38/Binding site: carbohydrate (Thr) (covalent) #status predicted
 F:178-179/Cleavage site: Lys-Ile (plasmin) #status experimental
 F:224, 275, 376/Active site: His, Asp, Ser #status experimental
 F:322/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 100.0%; Score 34; DB 1; Length 431;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHRGS 6
 DB 199 RHRGS 204

RESULT 2
 URBAY
 u-plasminogen activator (EC 3.4.21.73) precursor - yellow baboon
 C:Species: Papio cynocephalus, Papio hamadryas cynocephalus (yellow baboon)
 C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Jun-1999
 C:Accession: S14687; S086521
 R: Au, T.P.T.; Mang, T.W.; Clowes, A.W.
 Nucleic Acids Res. 18, 3411, 1990
 A:Title: Nucleotide and deduced amino acid sequences of baboon urokinase-type plasmin
 A:Reference number: S14687; MUID:90287734
 A:Accession: S14687
 A:Molecule type: mRNA
 A:Residues: 1-433 <AU>
 A:Cross-references: EMBL:X51935; NID:938130; PIDN:CAA36200.1; PID:q38131
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology;
 C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F:1-20/Domain: signal sequence #status predicted <Sig>
 F:21-176/Product: plasminogen activator chain A #status predicted <ACH>
 F:30-61/Domain: EGF homology <EGF>
 F:69-150/Domain: kringle homology <KR>
 F:178-433/Product: plasminogen activator chain B #status predicted <ECH>
 F:178-431/Domain: trypsin homology <TRY>
 F:167-298; 208-224; 216-287; 315-384; 347-363; 374-402/Disulfide bonds: #status predicted
 F:223, 274, 378/Active site: His, Asp, Ser #status predicted
 F:324/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 34; DB 1; Length 433;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHRGS 6
 DB 198 RHRGS 203

RESULT 3
 JN0560
 u-plasminogen activator (EC 3.4.21.73) precursor - bovine
 N:Alternate names: uPA
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
 C:Accession: JN0560
 R:Kraetzschmar, J.; Haendler, B.; Kojima, S.; Rifkin, D.B.; Schleuning, W.D.
 Gene 125, 177-183, 1993
 A:Title: Bovine urokinase-type plasminogen activator and its receptor: cloning and in
 A:Reference number: JN0560; MUID:93216119
 A:Accession: JN0560

A:Molecule type: mRNA
 A:Residues: 1-433 <KRA>
 A:Cross-references: GB:L03546; NID:g163800; PIDN:AA51419.1; PID:g163801
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F:1-20/Domain: signal sequence #status predicted <Sig>
 F:21-179/Product: plasminogen activator chain A #status predicted <MA1>
 F:21-179/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
 F:33-64/Domain: EGF homology <EGF>
 F:72-153/Domain: kringle homology <KR>
 F:181-433/Product: plasminogen activator chain B #status predicted <MA2>
 F:181-433/Product: trypsin homology <TRY>
 F:170-301,211-227,219-290,315-384,374-402/Disulfide bonds: #status predicted
 F:228,277,378/Active site: His, Asp, Ser #status predicted

Query Match 100.0%; Score 34; DB 1; Length 433;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRGS 6
 |||||
 Db 201 RHRGS 206

RESULT 4

U:plasmaogen activator (EC 3.4.21.73) precursor - p19
 N:Alternate names: UPA
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 04-Dec-1986 #sequence_revision 17-Mar-1987 #text_change 07-Aug-1998
 C:Accession: A00932
 R:Nagamine, Y.; Pearson, D.; Altus, M.S.; Reich, E.
 Nucleic Acids Res. 12, 9525-9541, 1984
 A>Title: CDNA and gene nucleotide sequence of porcine plasminogen activator.
 A:Reference number: A00932; MID:85087954
 A:Accession: A00932
 A:Molecule type: DNA
 A:Residues: 1-240, 'H', 242-442 <NMG1>
 A:Experimental source: kidney cell line LLC-PK1
 R:Nagamine, Y.
 Submitted to the Protein Sequence Database, December 1986
 A:Reference number: A37566
 A:Contents: annotation; correction to residue 241
 C:Genetics:
 A:introns: 19/3; 31/1; 67/1; 125/2; 165/1; 238/2; 288/1; 335/1; 384/3
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F:1-20/Domain: signal sequence #status predicted <Sig>
 F:21-188/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
 F:33-64/Domain: EGF homology <EGF>
 F:72-153/Domain: kringle homology <KR>
 F:190-442/Product: urokinase-type plasminogen activator chain B #status predicted <BGH>
 F:190-442/Product: trypsin homology <TRY>
 F:115/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:119-310,220-236,228-299,324-393,356-372,383-411/Disulfide bonds: #status predicted
 F:235,286,387/Active site: His, Asp, Ser #status predicted

Query Match 100.0%; Score 34; DB 1; Length 442;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRGS 6
 |||||
 Db 210 RHRGS 215

RESULT 5

adenosylhomocysteinease (EC 3.3.1.1) DKFp564A1523 - human (fragment)
 N:Alternate names: protein DKFp564A1523.1
 C:Species: Homo sapiens (man)

C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 23-Jul-1999
 C:Accession: T08681
 R:Hamut, R.; Heuber, D.; Wewes, H.W.; Gassenhber, J.; Wleemann, S.
 Submitted to the Protein Sequence Database, May 1999
 A:Reference number: Z16469
 A:Accession: T08681
 A:Molecule type: mRNA
 A:Residues: 1-597 <NMA>
 A:Cross-references: EMBL:AL049954
 A:Experimental source: fetal brain; clone DKFp564A1523
 C:Genetics:
 A:Note: DKFp564A1523.1
 C:Superfamily: adenosylhomocysteinease
 C:Keywords: NAD; thioether hydrolase
 F:379-408/Region: beta-alpha-beta NAD nucleotide-binding fold
 F:244,278/Active site: Cys #status predicted

Query Match 100.0%; Score 34; DB 2; Length 597;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRGS 6
 |||||
 Db 20 RHRGS 25

RESULT 6

S52714
 sericinB - silkworm
 C:Species: Bombyx mori (silkworm)
 C:Date: 19-May-1995 #sequence_revision 21-Jul-1995 #text_change 09-Sep-1997
 C:Accession: S52714
 R:Garel, A.A.; Deleage, G.G.; Prudhomme, J.J.
 Submitted to the EMBL Data Library, March 1995
 A:Description: Structure and organisation of the Bombyx mori sericin I gene and of t
 A:Reference number: S52714
 A:Accession: S52714
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1217 <GAR>
 A:Cross-references: EMBL:248602; NID:g755699; PID:g755700

Query Match 100.0%; Score 34; DB 2; Length 1217;
 Best Local Similarity 100.0%; Pred. No. 38;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRGS 6
 |||||
 Db 731 RHRGS 736

RESULT 7

T23024
 hypothetical protein H02112.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T23024
 R:McClay, K.
 Submitted to the EMBL Data Library, March 1997
 A:Reference number: Z19656
 A:Accession: T23024
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1332 <WIL>
 A:Cross-references: EMBL:292789; PIDN:CA07215.1; GSPDB:GN00022; CESP:H02112.1
 A:Experimental source: clone H02112
 C:Genetics:
 A:Gene: CESP:H02112.1
 A:Map position: 4
 A:introns: 12/3; 29/1; 60/3; 246/1; 449/1; 756/1; 849/1; 961/1; 1103/3; 1214/3

Query Match 100.0%; Score 34; DB 2; Length 1332;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRGS 6
|||||
DB 433 RHRGS 438

RESULT 8

S32789 Toxin II-14 - scorpion (Centruroides noxius)

C:Species: Centruroides noxius

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 07-May-1999

C:Accession: S32789

R:Possani, L.D.; Martin, B.M.; Svendsen, I.; Rode, G.S.; Erickson, B.W.

Biochem. J. 229, 739-750, 1985

A:Title: Scorpion toxins from Centruroides noxius and Tityus serrulatus. Primary structure

A:Reference number: S32789; MUID:86025386

A:Accession: S32789

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-64 <POS>

C:Superfamily: scorpion neurotoxin

Query Match

91.2%; Score 31; DB 2; Length 64;
Best Local Similarity 83.3%; Pred. No. 11;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRGS 6
|||||
DB 31 RHRGS 36

RESULT 9

JN0671 Na+-channel-blocking toxin (clone engtlv) precursor - scorpion (Centruroides noxius)

C:Species: Centruroides noxius

C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 16-Jul-1999

C:Accession: JN0671

R:Beccerilli, B.; Vazquez, A.; Garcia, C.; Corona, M.; Bolivar, F.; Possani, L.D.

Gene 128, 165-171, 1993

A:Title: Cloning and characterization of cDNAs that code for Na+-channel-blocking toxins

A:Reference number: JN0669; MUID:93292983

A:Accession: JN0671

A:Molecule type: mRNA

A:Residues: 1-86 <BEC>

A:Cross-references: GB:L05062; NID:9304570; PIDN:AAA28287.1; PID:9304571

A:Experimental source: venom gland

C:Superfamily: scorpion neurotoxin

C:Keywords: toxin

F:1-19/Domain: signal sequence #status predicted <STO>

F:20-86/Product: Na+-channel-blocking toxin (clone engtlv) #status predicted <MAT>

Query Match

91.2%; Score 31; DB 2; Length 86;
Best Local Similarity 83.3%; Pred. No. 15;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRGS 6
|||||
DB 50 RHRGS 55

RESULT 10

C84688 Probable C2H2-type zinc finger protein (imported) - Arabidopsis thaliana.

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: C84688

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Unayam, L.; Tallon, eus, D.; Nieman, M.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487

A:Accession: C84688

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-156 <STO>

A:Cross-references: GB:AE002093; NID:94580383; PIDN:AMD24362.1; GSPDB:GN00139

A:Gene: At2g28710

A:Map position: 2

Query Match 91.2%; Score 31; DB 2; Length 156;
Best Local Similarity 83.3%; Pred. No. 25;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRGS 6
|||||
DB 99 RHRGS 104

RESULT 11

E83606 hypothetical protein PA0310 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: E83606

R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;

Adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapid, K.; L

.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa

A:Reference number: A82950; MUID:20437337

A:Accession: E83606

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-232 <STO>

A:Cross-references: GB:AE004469; GB:AE004091; NID:9946152; PIDN:AA03699.1; GSPDB:GN

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA0310

Query Match 91.2%; Score 31; DB 2; Length 232;
Best Local Similarity 83.3%; Pred. No. 36;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRGS 6
|||||
DB 14 RHRGT 19

RESULT 12

T43147 hypothetical protein - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000

C:Accession: T43147

R:Yoshioke, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.

DNA Res. 4, 363-369, 1997

A:Title: Identification of open reading frames in Schizosaccharomyces pombe cDNAs.

A:Reference number: Z17323; MUID:98162722

A:Accession: T43147

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-310 <YOS>

A:Cross-references: EMBL:D89241; NID:91749689; PIDN:BA13902.1; PID:91749690

A:Experimental source: strain PR745

Query Match 91.2%; Score 31; DB 2; Length 310;
 Best Local Similarity 83.3%; Pred. No. 47;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRGS 6
 |||||
 Db 217 KHRGS 222

RESULT 13

G70800
 Probable ABC transporter ATP-binding protein - Mycobacterium tuberculosis (strain H37Rv)
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 17-Mar-2000
 C:Accession: G70800
 R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, A.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellern, S.; Squares, S.; Maitre 393; 337-344, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98293587
 A:Accession: G70800
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-376 <COL>
 A:Cross-references: GB:AL022121; GB:AL123456; NID:g3261559; PIDN:CA18080.1; PID:el26461
 A:Experimental source: strain H37Rv
 C:Genetics:
 A:Gene: prov
 C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
 C:Keywords: ATP
 E:18-212/Domain: ATP-binding cassette homology <ABC>

Query Match 91.2%; Score 31; DB 2; Length 376;
 Best Local Similarity 83.3%; Pred. No. 55;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRGS 6
 |||||
 Db 300 RHRGA 305

RESULT 14

JW0098
 carbazole dioxygenase (EC 1.14.-.-) terminal oxygenase component large chain carA - Spi
 C:Species: Sphingomonas sp.
 C:Date: 03-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 18-Jun-1999
 C:Accession: JW0098
 R:Shepherd, J.M.; Lloyd-Jones, G.
 Biochem. Biophys. Res. Commun. 247, 129-135, 1998
 A>Title: Novel carbazole degradation genes of Sphingomonas CB3: Sequence analysis, trans
 A:Reference number: JW0098; MUID:98300296
 A:Accession: JW0098
 A:Molecule type: mRNA
 A:Residues: 1-431 <SHE>
 A:Cross-references: GB:AF060489
 A:Experimental source: strain CB3
 C:Comment: This protein is the initial dioxygenase large (alpha) subunit in the carbazo
 C:Genetics:
 A:Gene: carA
 C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [2Fe
 C:Keywords: 2Fe-2S; aromatic hydrocarbon catabolism; metalloprotein; oxidoreductase; Rie
 E:65-113/Domain: Rieske [2Fe-2S] homology
 E:75,77,95,98/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status predi

Query Match 91.2%; Score 31; DB 2; Length 431;
 Best Local Similarity 83.3%; Pred. No. 62;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRGS 6

Db 76 RHRGA 81
 |||||

RESULT 15

T14339
 sucrose-proton transport protein - carrot
 N:Alternate names: sucrose/H⁺ symporter protein
 C:Species: Daucus carota (carrot)
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
 C:Accession: T14339
 R:Shakya, R.; Sturm, A.
 Plant Physiol. 118, 1473-1480, 1998
 A>Title: Characterization of source- and sink-specific sucrose/H⁺ symporters from ca
 A:Reference number: Z17991; MUID:99063785
 A:Accession: T14339
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-501 <SHA>
 A:Cross-references: EMBL:Y16766; NID:g296986; PIDN:CA176367.1; PID:g2969867
 A:Experimental source: cultivar Namtalse: leaf
 C:Genetics:
 A:Gene: SUT1a
 C:Superfamily: common tobacco sucrose transport protein

Query Match 91.2%; Score 31; DB 2; Length 501;
 Best Local Similarity 83.3%; Pred. No. 71;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRGS 6
 |||||
 Db 10 RHRGA 15

Search completed: March 18, 2002, 09:50:28
 Job time: 237 sec

Tue Mar 19 11:00:01 2002

us-09-544-665-1.rpt

Page 6

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: March 18, 2002, 09:47:31 ; Search time 24.53 Seconds

(without alignments)
8.968 Million cell updates/sec

Title: US-09-544-665-1
Perfect score: 34
Sequence: 1 RHRCGS 6

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	100.0	123	1 ANG2_BOVIN	P80928 bos taurus
2	34	100.0	431	1 UROK_HUMAN	P00749 homo sapien
3	34	100.0	433	1 UROK_BOVIN	Q05588 bos taurus
4	34	100.0	433	1 UROK_PAPCY	P06227 papio cynoc
5	34	100.0	442	1 UROK_PIG	P04185 sus scrofa
6	31	91.2	86	1 SCX1_CENNO	P15223 centruioid
7	31	91.2	86	1 SCX1_CENNO	P45665 centruioid
8	31	91.2	508	1 Y736_HAEIN	P44849 haemophilus
9	31	91.2	643	1 DNKA_STRE	Q9476 streptomyc
10	30	88.2	260	1 URK1_MOUSE	P55623 mus musculu
11	30	88.2	266	1 URK1_MOUSE	P21434 nicotiana d
12	30	88.2	271	1 RK2_SPIOL	P06508 spiracha ol
13	30	88.2	273	1 RK2_HORVU	P14096 hordeum vul
14	30	88.2	273	1 RK2_HAIZE	P17088 zea mays (m
15	30	88.2	273	1 RK2_ORYZA	P17351 oryza sativ
16	30	88.2	273	1 RK2_TREPA	Q81222 treponema p
17	30	88.2	274	1 RK2_ARATH	P56791 arabidopsis
18	30	88.2	274	1 RK2_EPTVI	P30065 epifagus vl
19	30	88.2	274	1 RK2_SIGNAL	P27107 sinapis alb
20	30	88.2	274	1 RK2_TOBAC	P06379 nicotiana t
21	30	88.2	275	1 RK2_CHLYU	P56367 chlorella v
22	30	88.2	275	1 RK2_ODOST	P49548 odontella s
23	30	88.2	275	1 RK2_ODOST	Q62954 picea abies
24	30	88.2	276	1 RK2_PICAB	Q62954 picea abies
25	30	88.2	276	1 RK2_PICAB	Q62954 picea abies
26	30	88.2	276	1 RK2_PICAB	Q62954 picea abies
27	30	88.2	276	1 RK2_PICAB	Q62954 picea abies
28	30	88.2	276	1 RK2_PICAB	Q62954 picea abies
29	30	88.2	276	1 RK2_PICAB	Q62954 picea abies
30	30	88.2	283	1 RK2_MYCGA	Q52335 mycoplasma
31	30	88.2	284	1 RK2_CHLYU	Q9P117 chlamydia m
32	30	88.2	284	1 RK2_CHLYU	Q9P117 chlamydia m
33	30	88.2	284	1 RK2_CHLYU	Q9P117 chlamydia m
				1 RL2_CHLTR	Q84530 chlamydia t

34	30	88.2	287	1 RL2_STNP6	Q24692 synechococ
35	30	88.2	295	1 RL2_AQUPE	Q92147 aquilex pyr
36	30	88.2	304	1 RL2_AQUPE	Q66434 aquilex aeo
37	30	88.2	342	1 HEAD_BPP21	P36270 bacterioph
38	30	88.2	455	1 PHR_STRGR	P12768 streptomyc
39	30	88.2	461	1 CDS1_RAT	Q35052 r phosphat
40	30	88.2	525	1 VCL_THRC	Q43358 theobroma c
41	30	88.2	791	1 K6PP_RABIT	P47859 oryctolagus
42	30	88.2	1096	1 PULA_KREAE	P07811 kiebsiella
43	30	88.2	1323	1 NME4_MOUSE	Q03391 mus musculu
44	30	88.2	1323	1 NME4_MOUSE	Q62645 rattus norv
45	30	88.2	1336	1 NME4_HUMAN	O15399 homo sapien

ALIGNMENTS

RESULT 1					
ANG2_BOVIN		STANDARD:	PRT:	123 AA.	
ID	ANG2_BOVIN				
AC	P80928				
DT	01-NOV-1997 (Rel. 35, Created)				
DT	01-NOV-1997 (Rel. 35, Last sequence update)				
DT	15-DEC-1998 (Rel. 37, Last annotation update)				
DE	ANGIOGENIN-2 (EC 3.1.27.-)				
GN	ANG2				
OS	Bos taurus (Bovine)				
OC	Eumariota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;				
OC	Bovidae; Bovine; Bos.				
OX	NCBI_TaxID=9913;				
RN	[1]				
RC	SEQUENCE.				
RC	TISSUE=Serum, and Milk;				
RX	MEDLINE=97409980; PubMed-9266695;				
RA	Strydom D.J., Bond M.D., Vallee B.L.;				
RT	"An angiogenic protein from bovine serum and milk -- purification and				
RT	primary structure of angiogenin-2.";				
RL	Eur. J. Biochem. 247:535-544(1997).				
CC	-1- FUNCTION: BINDS TIGHTLY TO PLACENTAL RIBONUCLEASE INHIBITOR AND				
CC	HAS VERY LOW RIBONUCLEASE ACTIVITY. HAS POTENT ANGIOGENIC				
CC	ACTIVITY. ANGIOGENIN INDUCES VASCULARIZATION OF NORMAL AND				
CC	MALIGNANT TISSUES. ABOLISHES PROTEIN SYNTHESIS BY SPECIFICALLY				
CC	HYDROLYZING CELLULAR TRNS.				
CC	-1- TISSUE SPECIFICITY: SERUM, AND MILK.				
CC	-1- SIMILARITY: BELONGS TO THE PANCREATIC RIBONUCLEASE FAMILY.				
DR	HSP: P10152; IGID.				
DR	Interpro: IPR001427; RNaseA.				
DR	Pfam: PF00074; RNaseA. 1.				
DR	PRINTS: PR00794; RIBONUCLEASE.				
DR	PRODOM: PD000535; RNaseA. 1.				
DR	SMART: SM00082; RNase_Pc. 1.				
DR	PROSITE: PS00177; RNase_PANCREATIC. 1				
KW	Hydrolase; Nuclease; Endonuclease; Vascularization;				
KW	Protein synthesis inhibitor; Glycoprotein;				
FT	MOD_RES 1				
FT	ACT_SITE 12 12				
FT	ACT_SITE 39 39				
FT	ACT_SITE 113 113				
FT	DISULFID 25 80				
FT	DISULFID 38 91				
FT	DISULFID 56 106				
FT	CARBOHYD 33				
SQ	SEQUENCE 123 AA; 14522 MW; B703B9839919ED2F CRC64;				

Query Match 100.0%; Score 34; DB 1; Length 123;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHRCGS 6
|||||
DB 81 RHRCGS 86

RESULT 2
UROK_HUMAN STANDARD: PRT: 431 AA.

ID UROK_HUMAN
AC P00749.
DT 21-JUL-1986 (Rel. 01, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE UROKINASE-TYPE PLASMINOGEN ACTIVATOR PRECURSOR (EC 3.4.21.73) (UPA)
DE (U-PLASMINOGEN ACTIVATOR).
GN PLAU.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85215647; PubMed=2987867;
RA Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Blas F.;
RT "The human urokinase-plasminogen activator gene and its promoter.";
RL Nucleic Acids Res. 13:2759-2771(1985).
RN [2]
RP SEQUENCE FROM N.A.
RA Holmes W.E., Pennica D., Blaber M., Rey M.W., Guenzler W.A.,
RA Steffens G.J., Heyneker H.L.;
RT "Cloning and expression of the gene for pro-urokinase in Escherichia
RT coli.";
RL Biotechnology 3:923-929(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=86056954; PubMed=2415429;
RA Nagai M., Hiramatsu R., Kaneda T., Hayasuke N., Atimura H.,
RA Nishida M., Suyama T.;
RT "Molecular cloning of cDNA coding for human preprourokinase.";
RL Gene 35:183-184(1985).
RN [4]
RP SEQUENCE OF 66-431 FROM N.A.
RX MEDLINE=84272706; PubMed=6589620;
RA Verde P., Stoppelli M.P., Galeffi P., di Nocera P., Blas F.;
RT "Identification and primary sequence of an unspliced human urokinase
RT poly(A)+ RNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984).
RN [5]
RP SEQUENCE OF 21-177.
RX MEDLINE=83055084; PubMed=6754569;
RA Guenzler W.A., Steffens G.J., Oetting F., Kim S.-M.A., Frankus E.,
RA Flohe L.;
RT "The primary structure of high molecular mass urokinase from human
RT urine. The complete amino acid sequence of the A chain.";
RL Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165(1982).
RN [6]
RP SEQUENCE OF 156-176 AND 179-224.
RX MEDLINE=83003608; PubMed=6744941;
RA Schaller J., Nick H., Rickli E.E., Gillesen D., Legler W.,
RA Studer R.O.;
RT "Human low-molecular-weight urinary urokinase. Partial
RT characterization and preliminary sequence data of the two polypeptide
RT chains.";
RL Eur. J. Biochem. 125:251-257(1982).
RN [7]
RP SEQUENCE OF 158-410.
RX MEDLINE=8305099; PubMed=6754572;
RA Steffens G.J., Guenzler W.A., Oetting F., Frankus E., Flohe L.;
RT "The complete amino acid sequence of low molecular mass urokinase
RT from human urine.";
RL Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1058(1982).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX MEDLINE=96000858; PubMed=8591045;
RA Spriggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D.,
RA Dobson C.M., Stuart D.I., Jones E.Y.;
RT "The crystal structure of the catalytic domain of human

RT urokinase-type plasminogen activator.";

RL Structure 3:681-691(1995).

RN [9]
RP STRUCTURE BY NMR.
RX MEDLINE=89127526; PubMed=2536903;
RA Oswald R.E., Boguski M.J., Bamberger M., Smith R.A.G., Dobson C.M.;
RT "Dynamics of the multidomain fibrinolytic protein urokinase from two-
RT dimensional NMR.";
RL Nature 337:579-582(1989).
RN [10]
RP STRUCTURE BY NMR OF 67-155.
RX MEDLINE=93003110; PubMed=1327118;
RA Li X., Smith R.A.G., Dobson C.M.;
RT "Sequential 1H NMR assignments and secondary structure of the kringle
RT domain from urokinase.";
RL Biochemistry 31:9562-9571(1992).
RN [11]
RP STRUCTURE BY NMR OF 67-155.
RX MEDLINE=94149701; PubMed=8107091;
RA Li X., Bokman A.M., Llinas M., Smith R.A.G., Dobson C.M.;
RT "Solution structure of the kringle domain from urokinase-type
RT plasminogen activator.";
RL J. Mol. Biol. 235:1548-1559(1994).
RN [12]
RP VARIANT LEU-141.
RX MEDLINE=96186279; PubMed=8652631;
RA Yoshimoto M., Ushiyama Y., Sakai M., Tamaki S., Hara H., Takahashi K.,
RA Swasak Y., Hanada K.;
RT "Characterization of single chain urokinase-type plasminogen
RT activator with a novel amino-acid substitution in the kringle
RT structure.";
RL Biochim. Biophys. Acta 1293:83-89(1996).
RN [13]
RP VARIANT LEU-141.
RX MEDLINE=97218551; PubMed=9065988;
RA Conne B., Bercezy M., Bellin D.;
RT "Detection of polymorphisms in the human urokinase-type plasminogen
RT activator gene.";
RL Thromb. Haemost. 77:434-435(1997).
RN [14]
RP VARIANT LEU-141.
RX MEDLINE=9737920; PubMed=9194591;
RA Turkmen B., Schmitt M., Schmalfeldt B., Trommler P., Hell W.,
RA Creutzburg S., Graeff H., Magdolen V.;
RT "Mutation analysis of the genes encoding urokinase-type plasminogen
RT activator (UPA) and its inhibitor PAI-1 in advanced ovarian cancer.";
RL Electrophoresis 18:686-689(1997).
RN [15]
RP FUNCTION: POTENT PLASMINOGEN ACTIVATOR AND IS CLINICALLY USED FOR
RP THERAPY OF THROMBOLYTIC DISORDERS.
CC -I- CATALYTIC ACTIVITY: SPECIFIC CLEAVAGE OF ARG-1-VAL BOND IN
CC PLASMINOGEN TO FORM PLASMIN.
CC -I- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW
CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN.
CC PHARMACEUTICAL: AVAILABLE UNDER THE NAME ABOKINASE (ABBOTT). USED
CC IN PULMONARY EMBOLISM (PE) TO INITIATES FIBRINOLYSIS.
CC -I- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -I- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPsin FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

```

CC -----
DR EMBL: X02419; CAA26268.1; -
DR EMBL: M15476; AAA61253.1; -
DR EMBL: D00244; BAA00175.1; -
DR EMBL: K03226; AAC97138.1; -
DR EMBL: K02286; AAA61252.1; -
DR EMBL: A21571; CAA01559.1; -
DR EMBL: A18397; CAA01390.1; -
DR PIR: A00931; UKHD; -
DR PIR: A32974; A32974; -
DR PDB: 1KDU; 31-OCT-93; -
DR PDB: 1LMW; 29-JAN-96; -
DR PDB: 1URK; 08-MAY-95; -
DR MEROPS: S01.231; -
DR GLYCOSEARCH: P00749; -
DR MIM: 191840; -
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Trypsin.
DR Pfam: PF00083; Kringle; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00130; KR; 1.
DR SMART: SM00020; TRYP_SPE; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; FALSE_NEG.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS00070; KRINGLE_2; 1.
DR PROSITE: PS0240; TRYPsin_DOM; 1.
DR PROSITE: PS00134; TRYPsin_HIS; 1.
DR PROSITE: PS00135; TRYPsin_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; glycoprotein;
KW Kringle; EGF-like domain; zymogen; Signal; Polymorphism;
KW Pharmaceutical; 3D-structure.
FT SIGNAL 1 20
FT CHAIN 21 431 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 177 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 156 177 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 179 431 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 179 431 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT DOMAIN 27 63 B CHAIN.
FT DOMAIN 70 151 EGF-LIKE.
FT DOMAIN 152 177 KRINGLE.
FT SITE 177 178 CONNECTING PEPTIDE.
FT DOMAIN 179 431 CLEAVAGE (DURING ZYMOGEN ACTIVATION).
FT DISULFID 31 39 BY SIMILARITY.
FT DISULFID 33 51 BY SIMILARITY.
FT DISULFID 33 62 BY SIMILARITY.
FT DISULFID 53 62 BY SIMILARITY.
FT DISULFID 168 299 INTERCHAIN.
FT DISULFID 209 225
FT DISULFID 217 288
FT DISULFID 313 382
FT DISULFID 345 361
FT DISULFID 372 400
Query Match 100.0%; Score 34; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.8; Mismatches 0; Gaps 0;
Matches 6; Conservative 0; Indels 0;

```

```

DT 01-FEB-1994 (rel. 28, last sequence update)
DT 01-FEB-1997 (rel. 35, last annotation update)
DE UROKINASE-TYPE PLASMINOGEN ACTIVATOR PRECURSOR (EC 3.4.21.73) (UPA)
DE (U-PLASMINOGEN ACTIVATOR).
DE PLAU.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Aortic endothelium.
RX MEDLINE=93216119; PubMed=8385052;
RA Kretzschmar J., Haendler B., Kojima S., Rifkin D.B.,
RA Schleuning W.-D.;
RT "Bovine urokinase-type plasminogen activator and its receptor:
RT cloning and induction by retinoic acid";
RL Gene 125:177-183(1993).
RN [2]
RP SEQUENCE OF 12-433 FROM N.A.
RC TISSUE-Kidney;
RA Ravn P., Berglund L., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen activators UPA
RT and tPA.";
RL Int. J. Cell. Physiol. 167:605-617(1995).
CC -1- CATALYTIC ACTIVITY: SPECIFIC CLEAVAGE OF ARG-1-VAL BOND IN
CC PLASMINOGEN TO FORM PLASMIN.
CC -1- INDUCTION: BY RETINOIC ACID.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
DR EMBL: L03546; AAA51419.1; -
DR EMBL: X85801; CAA59796.1; -
DR PIR: JN0560; JN0560.
DR HSSP: P00749; 1LMW.
DR MEROPS: S01.231; -
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Trypsin.
DR Pfam: PF00083; Kringle; 1.
DR Pfam: PF00051; Kringle; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00130; KR; 1.
DR SMART: SM00020; TRYP_SPE; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; FALSE_NEG.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS00070; KRINGLE_2; 1.
DR PROSITE: PS0240; TRYPsin_DOM; 1.
DR PROSITE: PS00134; TRYPsin_HIS; 1.
DR PROSITE: PS00135; TRYPsin_SER; 1.
KW plasminogen activation; Hydrolase; Serine protease; glycoprotein;
KW Kringle; EGF-like domain; Signal; zymogen.
FT SIGNAL 1 20
FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 179 CHAIN A (BY SIMILARITY).
FT CHAIN 181 433 CHAIN B (BY SIMILARITY).
FT DOMAIN 29 65 EGF-LIKE.
FT DOMAIN 72 153 KRINGLE.
FT DOMAIN 154 180 CONNECTING PEPTIDE.
FT DOMAIN 181 433 SERINE PROTEASE.

```

FT DISULFID 33 41 BY SIMILARITY.
 FT DISULFID 35 53 BY SIMILARITY.
 FT DISULFID 55 64 BY SIMILARITY.
 FT DISULFID 170 301 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 211 227 BY SIMILARITY.
 FT DISULFID 219 290 BY SIMILARITY.
 FT DISULFID 315 384 BY SIMILARITY.
 FT DISULFID 347 363 BY SIMILARITY.
 FT DISULFID 374 402 BY SIMILARITY.
 FT ACT_SITE 226 272 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 277 277 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 378 378 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT CONFLICT 189 189 A -> 1 (IN REF. 2).
 SO SEQUENCE 433 AA: 4DE1BBD4DA47027A CRC64;

Query Match 100.0%; Score 34; DB 1; Length 433;
 Best Local Similarity 100.0%; Pred. No. 5.9;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRRGS 6
 DB 201 RHRRGS 206

RESULT 4
 ID UROK_PAPCY STANDARD; PRT: 433 AA.
 AC P16227;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE UROKINASE-TYPE PLASMINOGEN ACTIVATOR PRECURSOR (EC 3.4.21.73) (UPA)
 DE (U-PLASMINOGEN ACTIVATOR).
 GN PLAU.
 OS Papio cynocephalus (Yellow baboon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Papio.
 OC NCBI_TaxID=9556;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=thoracic aorta;
 RC MEDLINE=90287734; PubMed=21133276;
 RA Au Y.F.T., Wang T.W., Clowes A.W.;
 RT Nucleotide and deduced amino acid sequences of baboon urokinase-type
 RL Nucleic Acids Res. 18:3411-3411(1990).
 CC -1- CATALYTIC ACTIVITY: SPECIFIC CLEAVAGE OF ARG-1-VAL BOND IN
 CC PLASMINOGEN TO FORM PLASMIN.
 CC -1- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
 CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
 CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW
 CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPsin FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X51935; CAA36200.1;
 DR PIR: S14687; UKBAY.
 DR HSSP: P00749; ILMW.
 DR MEROPS: S01.231;
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000561; EGF-like.

DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR001254; Trypsin.
 DR Pfam: PF00051; Kringle_1.
 DR Pfam: PF00089; Trypsin_1.
 DR PRINTS: PR00722; CHYMOTRYPsin.
 DR SMART: SM00181; EGF_1.
 DR SMART: SM00130; KR_1.
 DR SMART: SM00020; TRYP_Spc_1.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; FALSE NEG.
 DR PROSITE: PS00021; KRINGLE_1; 1.
 DR PROSITE: PS00070; KRINGLE_2; 1.
 DR PROSITE: PS00240; TRYPsin_DOM_1.
 DR PROSITE: PS00134; TRYPsin_HIS_1.
 DR PROSITE: PS00135; TRYPsin_SER_1.
 DR PLASMINOGEN activation; Hydrolyase; Serine protease; Glycoprotein;
 KW Kringle; EGF-like domain; Zymogen; Signal.
 FT SIGNAL 1 20
 FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
 FT CHAIN 21 176 CHAIN A (BY SIMILARITY).
 FT CHAIN 155 176 SHORT A CHAIN (A1) (BY SIMILARITY).
 FT CHAIN 178 433 CHAIN B (BY SIMILARITY).
 FT DOMAIN 26 62 EGF-LIKE.
 FT DOMAIN 69 150 KRINGLE.
 FT DOMAIN 151 177 CONNECTING PEPTIDE.
 FT DOMAIN 178 433 SERINE PROTEASE.
 FT DISULFID 30 38 BY SIMILARITY.
 FT DISULFID 32 50 BY SIMILARITY.
 FT DISULFID 52 61 BY SIMILARITY.
 FT DISULFID 167 298 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 208 224 BY SIMILARITY.
 FT DISULFID 216 287 BY SIMILARITY.
 FT DISULFID 315 384 BY SIMILARITY.
 FT DISULFID 347 363 BY SIMILARITY.
 FT DISULFID 374 402 BY SIMILARITY.
 FT ACT_SITE 223 223 CHARGE RELAY SYSTEM.
 FT ACT_SITE 274 274 CHARGE RELAY SYSTEM.
 FT ACT_SITE 378 378 CHARGE RELAY SYSTEM.
 FT CARBOHYD 324 324 N-LINKED (GLCNAC...), (BY SIMILARITY).
 SO SEQUENCE 433 AA: 48595 MW: 816D22DFD0C8792 CRC64;

Query Match 100.0%; Score 34; DB 1; Length 433;
 Best Local Similarity 100.0%; Pred. No. 5.9;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRRGS 6
 DB 198 RHRRGS 203

RESULT 5
 ID UROK_PIG STANDARD; PRT: 442 AA.
 AC P04185;
 DT 20-MAR-1987 (Rel. 04, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE UROKINASE-TYPE PLASMINOGEN ACTIVATOR PRECURSOR (EC 3.4.21.73) (UPA)
 DE (U-PLASMINOGEN ACTIVATOR).
 GN PLAU.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 OC NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RC MEDLINE=85087954; PubMed=6096832;
 RA Nagamine Y., Pearson D., Altus W.S., Reich E.;
 RT Nucleic Acids Res. 12:9525-9541(1984).
 RN [2]

QY 1 RHRGS 6
DB 50 KHRGS 55

RESULT 7

SCX7_CENNO STANDARD: PRT: 86 AA.
ID SCX7_CENNO
AC P45665;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE TOXIN CNGTIV PRECURSOR.
OS Centruroides noxius (Mexican scorpion).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Buthoidea; Buthidae; Centruroides.
OX NCBI_Taxid=6878;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Venom;
RX MEDLINE-9329283; PubMed-8390386;
RA Becerril B., Vazquez A., Garcia C., Corona M., Boliyar F.,
RA Pessan L.D.;
RT Cloning and characterization of cDNAs that code for Na(+)-channel-
blocking toxins of the scorpion Centruroides noxius Hoffmann.;
RL Gene 128:165-171(1993).
CC -1- FUNCTION: BINDS TO SODIUM CHANNELS AND INHIBITS THE INACTIVATION
OF THE ACTIVATED CHANNELS, THEREBY BLOCKING NEURONAL TRANSMISSION.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC BETA-TOXIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L05062; AAA28287.1; -
DR HSSP: P01494; ZSN3.
DR InterPro: IPR003614; Knc1.
DR InterPro: IPR002061; Scorpion_toxin.
DR Pfam: PF00537; toxin_3; 1.
DR PRINTS: PR00284; TOXIN.
DR PRINTS: PR00285; SCORPIONTOXIN.
DR ProDom: PD000908; Scorpion_toxin; 1.
DR SMART: SM00505; Knc1.1.
KW Neurotoxin; Sodium channel inhibitor; Amladation; Signal.
FT SIGNAL 1 19
FT CHAIN 20 84
FT DISULFID 30 83
FT DISULFID 34 59
FT DISULFID 43 64
FT DISULFID 47 66
FT MOD_RES 84 84
SQ SEQUENCE 86 AA: 9560 MM: 4D6BA41F905EBBD0 CMC64;
(POTENTIAL)
AMIDATION (G-85 PROVIDE AMIDE GROUP)

Query Match 91.28; Score 31; DB 1; Length 86;
Best Local Similarity 83.38; Pred. No. 4.8;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 RHRGS 6
DB 50 KHRGS 55

RESULT 8
Y736_HAEIN

ID Y736_HAEIN STANDARD: PRT: 508 AA.
AC P44849;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL SODIUM-DEPENDENT TRANSPORTER H10736.
GN H10736
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subphylum; Pasteurellales;
OC Haemophilus.
OX NCBI_Taxid=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RD / KW20 / ATCC 51907;
RX MEDLINE-9550630; PubMed-7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Gnehm C.L., McDonald J.L., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
influenzae Rd.";
RL Science 269:496-512(1995).
CC -1- FUNCTION: PUTATIVE SODIUM-DEPENDENT TRANSPORTER.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
FAMILY (SMF).
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U32757; AAC22395.1; -
DR TIGR: H10736; -
DR InterPro: IPR00175; Na_neurotran_symport.
DR Pfam: PF00209; SMF; 1.
DR PROSITE: PS00610; NA_NEUROTRAN_SYM_1; 1.
DR PROSITE: PS00754; NA_NEUROTRAN_SYM_2; FALSE_NEG.
DR PROSITE: PS0267; NA_NEUROTRAN_SYM_3; 1.
KW Hypothetical protein; Transport; Transmembrane; Symport;
KW Complete proteome.
FT TRANSMEM 19 39
FT TRANSMEM 47 67
FT TRANSMEM 92 112
FT TRANSMEM 133 173
FT TRANSMEM 186 206
FT TRANSMEM 228 248
FT TRANSMEM 272 292
FT TRANSMEM 325 345
FT TRANSMEM 365 385
FT TRANSMEM 394 414
FT TRANSMEM 432 452
FT TRANSMEM 469 489
SQ SEQUENCE 508 AA: 55177 MM: 5CC5052D6E854BB9 CRC64;
(POTENTIAL)

Query Match 91.28; Score 31; DB 1; Length 508;
Best Local Similarity 83.38; Pred. No. 30;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 RHRGS 6
DB 74 RHRGA 79

```

RESULT 9
ID DNA_STRE STANDARD; PRT; 643 AA.
AC 092H76;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE CHROMOSOMAL REPLICATION INITIATOR PROTEIN DNA.
GN DNA.
OS Streptomyces reticuli.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycetaceae; Streptomyces.
OX NCBI_Taxid-1926;
RN [1]
RP SEQUENCE FROM N.A.
RC SPRAIN-TU45;
RA Majka J., Jakimowicz D., Messer W., Schremf H., Lisowski M.,
RA Zakrzewska-Czerwikowska J.,
RT Interactions of the Streptomyces lividans Initiator protein DnaA with
RT its target.
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE INITIATION AND REGULATION
CC OF CHROMOSOMAL REPLICATION. BINDS TO THE ORIGIN OF REPLICATION;
CC IT BINDS SPECIFICALLY DOUBLE-STRANDED DNA AT A 9 BP CONSENSUS
CC (DNA BOX): 5'-TTATC(C/A)A(C/A)-3'. DNAA BINDS TO ATP AND TO
CC ACIDIC PHOSPHOLIPIDS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE DNAA FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: AF07023; AAD08806.1;
CC InterPro: IPR003593; AAA.
CC InterPro: IPR001957; Bac_DnaA.
CC Pfam: PF00308; bac_dnaa; 1.
CC PRINTS: PRO0051; DNAA.
CC SMART: SM00382; AAA; 1.
CC PROSITE: PS01008; DNAA; 1.
CC DR DNA replication; DNA-binding; ATP-binding.
CC KW NP_BIND 344 351 ATP (POTENTIAL).
CC FT SEQUENCE 643 AA; 71317 MW; DB9E173DF2475B85 CRC64;
CC SQ
-----
Query Match 91.2%; Score 31; DB 1; Length 643;
Best Local Similarity 83.3%; Pred. No. 38;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 RHRGGS 6
DB 125 RHRGGA 130
-----
RESULT 10
ID URK1_MOUSE STANDARD; PRT; 260 AA.
AC P52623;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE URIDINE KINASE (EC 2.7.1.48) (URIDINE MONOPHOSPHOKINASE) (FRAGMENT).
GN UMPK.
OS Mus musculus (Mouse).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid-10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Brain;

```

```

RX MEDLINE-97108719; PubMed-8951040;
RA Ropp P.A., Traut T.W.;
RT "Cloning and expression of a cDNA encoding uridine kinase from mouse
RT brain."
RL Arch. Biochem. Biophys. 336:105-112(1996).
CC -1- CATALYTIC ACTIVITY: ATP + URIDINE -> ADP + UMP.
CC -1- PATHWAY: PYRIMIDINE SALVAGE PATHWAY.
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: L31783; AAB50568.1;
CC MGP: MGI:98904; Umpk.
CC InterPro: IPR001324; PRK.
CC Pfam: PF00485; PRK; 1.
CC TRANSFERASE; Kinase.
CC KW NON_TER 1
CC FT SEQUENCE 260 AA; 29622 MW; 080F1C82C59C80AC CRC64;
CC SQ
-----
Query Match 88.2%; Score 30; DB 1; Length 260;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RHRGG 5
DB 221 RHRGG 225
-----
RESULT 11
ID RK2_NICDE STANDARD; PRT; 266 AA.
AC P21434;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CHLOROPLAST 50S RIBOSOMAL PROTEIN L2.
GN RPL2.
OS Nicotiana debneyi (Debney's tobacco).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_Taxid-4089;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-84297246; PubMed-6089120;
RA Zurawski G., Bottomley W., Whitfield P.R.;
RT "Junctions of the large single copy region and the inverted repeats
RT in Spinacia oleracea and Nicotiana debneyi chloroplast DNA: sequence
RT of the genes for rRNAs and the ribosomal proteins S19 and L2."
RL Nucleic Acids Res. 12:6547-6558(1984).
CC -1- SIMILARITY: BELONGS TO THE L2P FAMILY OF RIBOSOMAL PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: X00798; CAB52367.1;
CC PIR: S07356; RSN12D.
CC Mendel: 5309; NICDE; rpl2.1.
CC InterPro: IPR002171; Ribosomal_L2.
CC Pfam: PF00181; Ribosomal_L2; 1.

```

DR PROSITE: PS00467; RIBOSOMAL.L2; 1.
 KW Ribosomal protein; Chloroplast.
 SO SEQUENCE 266 AA; 28901 MW; 8F4AAE3CE2ADF314 CRC64;

Query Match
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHRGG 5
 |||||
 Db 50 RHRGG 54

RESULT 12
 RK2_SPIOI STANDARD: PRT: 271 AA.
 AC P06509: O9THV0: Q9M311:
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DE 20-AUG-2001 (Rel. 40, Last annotation update)
 DE CHLOROPLAST 50S RIBOSOMAL PROTEIN L2 (RIBOSOMAL PROTEIN CS-L4).
 GN RPL2.
 OS Spinacia oleracea (Spinach).
 OC Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
 OX NCBI_TaxID=3562;
 RN [1]
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 1-10.
 RX MEDLINE=84297246; PubMed=6089120.
 RA Zurawski G., Bottomley W., Whitfield P.R.;
 RT "Junctions of the large single copy region and the inverted repeats
 in Spinacia oleracea and Nicotiana debneyi chloroplast DNA: sequence
 of the genes for tRNAHis and the ribosomal proteins S19 and L2.";
 RL Nucleic Acids Res. 12:6547-6558(1984).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Leaf;
 RA Yamauchi K., Knoblauch K., Subramanian A.R.;
 RT "Complete identification of ribosomal proteins in the large subunit of
 an organelle (chloroplast) ribosome.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. GEANT D'HIVER. AND CV. MONATOL;
 RX MEDLINE=21187424; PubMed=11292076;
 RA Schmitz-Linneweber C., Maier R.M., Alcaraz J.-P., Cottet A.,
 RA Herrmann R.G., Mache R.;
 RT "The plastid chromosome of spinach (Spinacia oleracea): complete
 nucleotide sequence and gene organization.";
 RL Plant Mol. Biol. 45:307-315(2001).
 RN [4]
 RP SEQUENCE OF 1-19 FROM N.A.
 RX MEDLINE=88203193; PubMed=3362671;
 RA Thomas F., Massenet O., Dorne A.H., Briat J.F., Mache R.;
 RT "Expression of the rpl23, rpl2 and rps19 genes in spinach
 chloroplasts.";
 RL Nucleic Acids Res. 16:2461-2472(1988).
 RN [5]
 RP METHYLATION, AND PARTIAL SEQUENCE.
 RC STRAIN=CV. ALVARO;
 RA Kamp R.M., Srinivasa B.R., von Knoblauch K., Subramanian A.R.;
 RT "Occurrence of a methylated protein in chloroplast ribosomes.";
 RL Biochemistry 26:5866-5870(1987).
 CC -1- SIMILARITY: BELONGS TO THE L2P FAMILY OF RIBOSOMAL PROTEINS.
 CC -1- CAUTION: REP.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
 CC FRAMESHIFT IN POSITION 227.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.ch/announce/>
 CC or send an email to license@isb.ch).

DR EMBL: AJ244023; CAB56543.3;
 DR EMBL: AJ400848; CAB88803.1;
 DR EMBL: X07462; CA30345.1;
 DR PIR: S07918; R5SP2.
 DR Mendel, 4699; SPIOI.rpl2.1.
 DR InterPro: IPR002171; Ribosomal.L2.
 DR Pfam: PF00181; Ribosomal.L2; 1.
 DR PROSITE: PS00467; RIBOSOMAL.L2; 1.
 KW Ribosomal protein; Chloroplast; Methylation.
 FT INT_MET 0
 FT MOD_RES 1
 FT SEQUENCE 271 AA; 29661 MW; B8E06C44FD93B63 CRC64;

Query Match
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHRGG 5
 |||||
 Db 47 RHRGG 51

RESULT 13
 RK2_HORVU STANDARD: PRT: 273 AA.
 AC P41086;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE CHLOROPLAST 50S RIBOSOMAL PROTEIN L2.
 GN RPL2.
 OS Hordeum vulgare (Barley).
 OC Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticeae; Hordeum.
 OX NCBI_TaxID=4513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. HAISA;
 RX MEDLINE=95086380; PubMed=7994178;
 RA Hees W.R., Hoch B., Zeltz P., Huebschmann T., Koessel H.,
 RA Boerner T.;
 RT "Inefficient rpl2 splicing in barley mutants with ribosome-deficient
 plastids.";
 RL Plant Cell 6:1455-1465(1994).
 CC -1- SIMILARITY: BELONGS TO THE L2P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.ch/announce/>
 CC or send an email to license@isb.ch).

DR EMBL: X78185; CA55028.1; ALT_SEQ.
 DR Mendel; 4044; HORVU.rpl2.1.
 DR InterPro: IPR002171; Ribosomal.L2.
 DR Pfam: PF00181; Ribosomal.L2; 1.
 DR PROSITE: PS00467; RIBOSOMAL.L2; 1.
 KW Ribosomal protein; Chloroplast.
 FT SEQUENCE 273 AA; 30045 MW; 4C204966ACB7DBE9 CRC64;

Query Match
 Best Local Similarity 100.0%; Pred. No. 26;
 Score 30; DB 1; Length 273;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRCG 5
|||||
Db 50 RHRCG 54

RESULT 14

NR2_MAIZE STANDARD; PRT: 273 AA.
AC P17788;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE CHLOROPLAST 50S RIBOSOMAL PROTEIN L2.
GN RPL2
OS Zea mays (Maize).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoidae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. PRGCMSSR37;
RX MEDLINE=90332419; PubMed=2377464;
RA KAVOUSI M., GLESE K., LATTINIA I.M., SUBRAMANIAN A.R.;
RT Nucleotide sequence and map positions of the duplicated gene for
RT maize (Zea mays) chloroplast ribosomal protein L2.
RL Nucleic Acids Res. 18:4244-4244(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95395841; PubMed=7666415;
RA MAIER R.M., NECKERMANN K., IGLOI G.L., KOESSEL H.;
RT The complete sequence of the maize chloroplast genome: gene content,
RT hotspots of divergence and fine tuning of genetic information by
RT transcript editing.
RL J Mol. Biol. 251:614-628(1995).
RN [3]
RP RNA EDITING OF INITIATOR CODON.
RX MEDLINE=91367263; PubMed=1653905;
RA HOCH B., MAIER R.M., APPEL K., IGLOI G.L., KOESSEL H.;
RT Editing of a chloroplast mRNA by creation of an initiation codon.
RL Nature 353:178-180(1991).
CC -1 SIMILARITY: BELONGS TO THE L2P FAMILY OF RIBOSOMAL PROTEINS.
CC -1 CAUTION: THE INITIATOR METHIONINE IS CREATED BY RNA EDITING.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X53066; CAA37241.1; -
DR EMBL: X86563; CAA60329.1; ALT_SEQ.
DR EMBL: X86563; CAA60371.1; -
DR EMBL: X62070; CAA3983.1; ALT_SEQ.
DR PIR: S10500; R52M2.
DR PIR: S10500; R52M2.
DR PIR: S10500; R52M2.
DR MaizeDB: 66413; -
DR MaizeDB: 11440; ZEMa:RPL2.1.
DR InterPro: IPR002171; Ribosomal_L2.
DR Pfam: PF00181; Ribosomal_L2; 1.
DR PROSITE: PS00467; RIBOSOMAL_L2; 1.
KW Ribosomal protein; Chloroplast; RNA editing.
SO SEQUENCE 273 AA; 30065 MW; BA65197231EA3CA0 CRC64;

Query Match 88.2%; Score 30; DB 1; Length 273;

Best Local Similarity 100.0%; Pred. No. 26; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRCG 5
|||||
Db 50 RHRCG 54

RESULT 15

NR2_ORYSA STANDARD; PRT: 273 AA.
AC P17351; P92327;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CHLOROPLAST 50S RIBOSOMAL PROTEIN L2.
GN RPL2-A AND RPL2-B.
OS Oryza sativa (Rice).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. NIPPONBARE;
RA Sugiyama M.;
RL Submitted (JUL-1989) to the EMBL/GenBank/DBJ databases.
RN [2]
RP COMPLETE GENOME.
RX MEDLINE=89364688; PubMed=2770692;
RA Hiratsuka J., Shimada H., Whittier R., Ishibashi T., Sakamoto M.,
RA Mori M., Kondo C., Honji Y., Sun C.-R., Meng B.-Y., Li Y.-Q.,
RA Kanno A., Nishizawa Y., Hirai A., Shinozaki K., Sugiyama M.;
RT The complete sequence of the rice (Oryza sativa) chloroplast genome:
RT intermolecular recombination between distinct tRNA genes accounts for
RT a major plastid DNA inversion during the evolution of the cereals.
RL Mol. Gen. Genet. 217:185-194(1989).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=89196901; PubMed=3240862;
RA Moon E., Wu R.;
RT Organization and nucleotide sequence of genes at both junctions
RT between the two inverted repeats and the large single-copy region in
RT the rice chloroplast genome.
RL Gene 70:1-12(1988).
CC -1 SIMILARITY: BELONGS TO THE L2P FAMILY OF RIBOSOMAL PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X15901; CAA33924.1; ALT_SEQ.
DR EMBL: X15901; CAA33928.1; ALT_SEQ.
DR EMBL: M2826; CAB25243.1; ALT_SEQ.
DR EMBL: L40578; AAD15254.1; ALT_SEQ.
DR PIR: J00270; R5R22.
DR PIR: J00270; R5R22.
DR Mendel: 5002; ORYSA:RPL2.1.
DR InterPro: IPR002171; Ribosomal_L2.
DR Pfam: PF00181; Ribosomal_L2; 1.
DR PROSITE: PS00467; RIBOSOMAL_L2; 1.
KW Ribosomal protein; Chloroplast.
FT CONFLICT 69 69- Q -> E (IN REF. 3).
FT CONFLICT 200 200 Q -> K (IN REF. 3).
FT CONFLICT 230 246 HGCGCGKAPRGRKPTT -> GAVGKGRPLVERKNPP
FT CONFLICT 230 246 (IN REF. 3).
SO SEQUENCE 273 AA; 29942 MW; AA41536C2AA8729B CRC64;

Query Match 88.2%; Score 30; DB 1; Length 273;

Best Local Similarity 100.0%; Pred. No. 26; Mismatches 0; Indels 0; Gaps 0;

Tue Mar 19 11:00:01 2002

us-09-544-665-1.rsp

Page 10

Matches	5;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Oy	1	RHRG	5						
Db	50	RHRG	54						

Search completed: March 18, 2002, 09:52:24
Job time: 293 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: March 18, 2002, 09:47:11 ; Search time 78.82 seconds
(without alignments)
11.135 Million cell updates/sec

Title: US-09-544-665-1
Perfect score: 34
Sequence: 1 RHRGS 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_17:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.potent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	100.0	128	6	097587
2	34	100.0	214	6	09XT70
3	34	100.0	411	4	015844
4	34	100.0	431	4	016618
5	34	100.0	444	13	09PTN4
6	34	100.0	597	4	09UG84
7	34	100.0	1217	5	017240
8	34	100.0	1332	5	045599
9	34	100.0	2146	13	09PUM3
10	34	100.0	2157	13	09PUM4
11	34	100.0	2171	13	09PUM5
12	34	100.0	2182	13	09PUM6
13	34	100.0	2321	13	09PUM7
14	34	100.0	2331	13	09PUM8
15	34	100.0	2332	13	09PUM9
16	34	100.0	2346	13	09PUM10
17	34	100.0	2357	13	09PUM11
18	34	100.0	4283	11	09ERV0
19	34	100.0	4293	11	008852

20	31	91.2	156	10	09S197	09s197 arabidopsis
21	31	91.2	232	2	09I611	09i611 pseudomonas
22	31	91.2	262	8	09XW04	09xw04 tetrahymena
23	31	91.2	297	2	09Z1L2	09z1l2 myxococcus
24	31	91.2	341	2	09F201	09f2q1 streptomyces
25	31	91.2	376	2	069724	069724 mycobacterium
26	31	91.2	431	2	085283	085283 sphingomonas
27	31	91.2	501	10	065929	065929 daucus carota
28	31	91.2	508	5	09W228	09w228 drosophila
29	31	91.2	675	2	069558	069558 mycobacterium
30	31	91.2	971	5	09V6X6	09v6v6 drosophila
31	30	88.2	54	10	041927	041927 arabidopsis
32	30	88.2	70	12	09Q0D0	09q0d0 herpes simp
33	30	88.2	70	12	09Q0C9	09q0c9 herpes simp
34	30	88.2	93	11	063730	063730 rattus norv
35	30	88.2	111	3	09P663	09p663 neurexina
36	30	88.2	118	2	09X035	09x035 thermotoga
37	30	88.2	130	1	09Y2E5	09y2e5 aeropyrum p
38	30	88.2	141	5	09V6R2	09v6r2 drosophila
39	30	88.2	145	5	09GWM2	09gwm2 leishmania
40	30	88.2	163	2	09R190	09r190 streptomyces
41	30	88.2	179	2	034419	034419 setaria ma
42	30	88.2	199	8	09GFC5	09g1j5 calycanthus
43	30	88.2	200	8	09GFC6	09g1j6 calycanthus
44	30	88.2	201	8	09MS05	09msq5 zambia furu
45	30	88.2	201	8	09MSP3	09msp3 nympheae od

ALIGNMENTS

RESULT	ID	097587	PRELIMINARY:	PRT:	128 AA.
AC	097587	01-MAY-1999 (TREMBLrel. 10, Created)			
DT	01-MAY-1999	(TREMBLrel. 10, Last sequence update)			
DT	01-JUN-2001	(TREMBLrel. 17, Last annotation update)			
DE	OROKINASE (FRAGMENT)				
OS	Oryctolagus cuniculus (Rabbit)				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.				
OX	NCBI_TaxID=9986;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-NEW ZEALAND WHITE;				
RX	MEDLINE=99057575; PubMed=9837780;				
RA	Reno C., Boykiv R., Martinez M.L., Hart D.A.;				
RT	"Temporal alterations in mRNA levels for proteinases and inhibitors				
RT	and their potential regulators in the healing medial collateral				
RT	ligament."				
CC	Biochem. Biophys. Res. Commun. 252:757-763(1998).				
CC	-I- SIMILARITY: TO SERINE PROTEASES, TRYPSIN FAMILY.				
CC	-I- SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).				
CC	EMBL; AF069711; AAC9503.1; -				
DR	InterPro: IPR001314; Chymotrypsin.				
DR	InterPro: IPR001254; Trypsin.				
DR	Pfam: PF00089; trypsin; 1.				
DR	PRINTS: PR00722; CHYMOTRYPSIN.				
DR	SMART: SM00020; TRYP-Spc; 1.				
DR	PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.				
KW	Hydrolase; Kinase; Serine protease.				
FT	NON-TER				
FT	NON-TER				
FT	NON-TER				
SEQUENCE	128 AA; 14328 MM; 1BC7ED30E71A06D CRC64;				

Query Match: 100.0%; Score 34; DB 6; Length 128;
Best Local Similarity: 100.0%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RHRGS 6 |||||

Db 8 RHRGS 13

RESULT 2

ID 09XT70 PRELIMINARY: PRT: 214 AA.

AC 09XT70: PRELIMINARY: PRT: 214 AA.
 DT 01-NOV-1999 (TREMblrel. 12, Created)
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE UROKINASE-TYPE PLASMINOGEN ACTIVATOR (FRAGMENT).
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LUNG;
 RA yin J. Igel) S;
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO SERINE PROTEASES, TRYPSIN FAMILY.
 CC -1- SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).
 DR EMBL: AF097647; AAD39351.1;
 DR HSSP: P00749; IKDU.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000001; Kringie.
 DR InterPro: IPR001254; Trypsin.
 DR Pfam: PF00051; Kringie_1.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00130; KR. 1.
 DR SMART: SM00020; TRYP_SPE. 1.
 DR PROSITE: PS00021; KRINGLE_1; 1.
 DR PROSITE: PS0070; KRINGLE_2; 1.
 DR PROSITE: PS00134; TRYPsin_HIS; UNKNOWN_1.
 DR KAM Hydrolyase: Kinase; Serine protease.
 FT NON_TER 214
 FT SEQUENCE 214 AA; 24314 MW; 69975C41C32B0D7E CRC64;
 SO

Query Match 100.0%; Score 34; DB 6; Length 214;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHRGS 6

Db 112 RHRGS 117

RESULT 3
 ID 015844 PRELIMINARY: PRT: 411 AA.

AC 015844: PRELIMINARY: PRT: 411 AA.
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE UROKINASE-TYPE PLASMINOGEN ACTIVATOR (FRAGMENT).
 GN UPA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=86056954; PubMed=2415429;
 RA Nagai M., Hiramatsu R., Kaneda T., Hayasuke N., Atimura H.,
 RA Nishida M., Suyama T.;
 RT "Molecular cloning of cDNA coding for human preprourokinase";
 RL Gene 36:183-188(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=96186279; PubMed=8652631;
 RX

RA Yoshimoto M., Ushiyama Y., Sakai M., Tamaki S., Hara H., Takahashi K.,

Sawasaki Y., Hanada K.;

RT "Characterization of single chain urokinase-type plasminogen activator

with a novel amino-acid substitution in the kringle structure.";

RL Biochim. Biophys. Acta 1293:83-89(1996).

CC -1- SIMILARITY: TO SERINE PROTEASES, TRYPSIN FAMILY.

CC -1- SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).

DR EMBL: D11143; BAA01919.1;
 DR HSSP: P00749; IURK.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR000001; Kringie.
 DR InterPro: IPR001254; Trypsin.
 DR Pfam: PF00051; Kringie_1; 1.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00018; Kringie.
 DR SMART: SM00181; EGF. 1.
 DR SMART: SM00130; KR. 1.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS00021; KRINGLE_1; 1.
 DR PROSITE: PS0070; KRINGLE_2; 1.
 DR PROSITE: PS00134; TRYPsin_HIS; 1.
 DR PROSITE: PS00135; TRYPsin_SER; 1.
 DR KAM Hydrolyase: Serine protease.
 FT NON_TER 1
 FT CONFICT 121 121 L -> P (IN REF. 1).
 FT SEQUENCE 411 AA; 46383 MW; F99AD080FD43FC96 CRC64;
 SO

Query Match 100.0%; Score 34; DB 4; Length 411;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHRGS 6

Db 179 RHRGS 184

RESULT 4
 ID 016618 PRELIMINARY: PRT: 431 AA.

AC 016618: PRELIMINARY: PRT: 431 AA.
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE UROKINASE PRECURSOR (EC 3.4.99.26).
 GN PLAU.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=85203359; PubMed=3888571;
 RA Jacobs F., Crivador A., Lortiau R., Brockly F., Colau B., Chuchana P.,
 RA van Elsen A., Herzog A., Boilen A.;
 RT "Molecular cloning, sequencing, and expression in Escherichia coli of human preprourokinase cDNA";
 RL DNA 4:139-146(1985).
 CC -1- SIMILARITY: TO SERINE PROTEASES, TRYPSIN FAMILY.
 CC -1- SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).
 DR EMBL: X02760; CAA26535.1;
 DR HSSP: P00749; IURK.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR000001; Kringie.
 DR InterPro: IPR001254; Trypsin.
 DR Pfam: PF00051; Kringie_1; 1.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00181; EGF. 1.

DR SMART: SM00130; KR: 1.
 DR SMART: SM00020; TRYP_SPC: 1.
 DR PROSITE: PS00022; EGF_1: 1.
 DR PROSITE: PS00021; KRINGLE_1: 1.
 DR PROSITE: PS00770; KRINGLE_2: 1.
 DR PROSITE: PS00134; TRYPsin_HIS: 1.
 DR PROSITE: PS00135; TRYPsin_SER: 1.
 KW Hydrolyase; Serine protease; Signal.
 FT SIGNAL
 FT CHAIN 21 431
 FT NON_TER 1
 SO SEQUENCE 431 AA; 48664 MW; 16082504B57AC18 CRC64;

Query Match 100.0%; Score 34; DB 4; Length 431;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRRGS 6
 DB 199 RHRRGS 204

RESULT 5
 O9PTN4 PRELIMINARY; PRT: 444 AA.
 AC O9PTN4;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMblrel. 13, Last annotation update)
 DE N-TYPE CALCIUM CHANNEL, ALPHA-1B SUBUNIT (FRAGMENT).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schitt M., Siderovski D., Jordan J.D., Brothers G., Snow B., Dho S.,
 RA Wolting C., McGlade J., De Vries L., Ortiz D., Diverse-Pierluissi M.,
 RT "GABA B receptor-induced N-type calcium channel modulation is mediated
 RT by interactions of the PDB domain of RGS12 with the tyrosine-
 RT phosphorylated channel.";
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF155829; AAF22237.1; -
 FT NON_TER 1
 FT SEQUENCE 444 AA; 50547 MW; 374D9E33FD27C56B CRC64;

Query Match 100.0%; Score 34; DB 13; Length 444;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRRGS 6
 DB 248 RHRRGS 253

RESULT 6
 O9UG84 PRELIMINARY; PRT: 597 AA.
 AC O9UG84;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE HYPOTHETICAL 66.7 KDA PROTEIN (FRAGMENT).
 GN DKF2P564A1523.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE-BRAIN;
 RA Wambutt R., Heubner D., Mewes H.M., Gassenhuber J., Miemann S.;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL049954; CAB43223.1; -
 DR HSSP: P10760; 1B3R.
 DR InterPro: IPR000043; Ado_hcyase.
 DR InterPro: IPR001220; Lectin_legb.
 DR Pfam: PF00670; AdoHcyase; 1.
 DR PROSITE: PS00738; ADOHCYASE_1; 1.
 DR PROSITE: PS00739; ADOHCYASE_2; UNKNOWN_1.
 DR PROSITE: PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
 KW Hypothetical protein.
 FT NON_TER 1
 SO SEQUENCE 597 AA; 66668 MW; 92EA6896796E501D CRC64;

Query Match 100.0%; Score 34; DB 4; Length 597;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRRGS 6
 DB 20 RHRRGS 25

RESULT 7
 O17240 PRELIMINARY; PRT: 1217 AA.
 AC O17240;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-NOV-1996 (TREMblrel. 01, Last annotation update)
 DE SERICINIB.
 OS Bombyx mori (Silk moth).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Bombycoidea; Bombycidae; Bombyx.
 OC NCBI_TaxID=7091;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-EUROPEAN 200X300; TISSUE-MEDIAN SILK GLAND;
 RA Gareil A.A., Deleage G.G., Prudhomme J.J.;
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Z48802; CAA88741.1; -
 SO SEQUENCE 1217 AA; 123410 MW; 4C5789F87F6866D0 CRC64;

Query Match 100.0%; Score 34; DB 5; Length 1217;
 Best Local Similarity 100.0%; Pred. No. 99;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRRGS 6
 DB 731 RHRRGS 736

RESULT 8
 O45599 PRELIMINARY; PRT: 1332 AA.
 AC O45599;
 DT 01-JUN-1998 (TREMblrel. 06, Created)
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE H02112.1 PROTEIN.
 GN H02112.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Peloderae; Caenorhabditis.
 OC NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA McIay K.;
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.

[2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE-94150718; PubMed-7906398;
 RA Milson R., Alnsough R., Anderson K., Baynes C., Berke M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favallo A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rinken L., Roopra A., Saunders D., Showkeen R.,
 RA Smailton N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thelery-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstein L., Wilkinson-Spratt J., Wohlman P.,
 RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RL Nature 368:32-38(1994).
 DR EMBL:292789; CAB07215.1;
 DR InterPro: IPR002357; Chitin_binding.
 DR Pfam: PF01607; Chitin_bind_2; 9.
 DR SMART: SM00494; ChEBD2; 12.
 SQ SEQUENCE 1332 AA; 145729 MW; 3D343EA7643131B6 CRC64;

Query Match 100.0%; Score 34; DB 5; Length 1332;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRGS 6
 DB 433 RHRGS 438

RESULT 9
 O9PUM3 PRELIMINARY; PRT; 2146 AA.
 ID O9PUM3;
 AC O9PUM3;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE N-TYPE CALCIUM CHANNEL ALPHA-1B CDB4 VARIANT.
 GN CACNA1B.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID-9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-DORSAL ROOT GANGLION NEURONS;
 RX MEDLINE-20044702; PubMed-10574919;
 RA Lu Q., Dunlap K.;
 RT "Cloning and Functional Expression of Novel N-type Ca2+ Channel
 RT Variants.";
 RL J. Biol. Chem. 274:34566-34575(1999).
 DR EMBL: AF173015; AAD51818.1;
 DR InterPro: IPR000636; Cation_chan_non_119.
 DR InterPro: IPR001682; Channel_pore_Ca_Na.
 DR InterPro: IPR001696; Na_channel.
 DR InterPro: IPR002077; Ca_channel.
 DR InterPro: IPR002111; Cat_channel_TrpL.
 DR Pfam: PF00520; Ion_trans_4.
 DR PRINTS: PR00167; CACHANNEL.
 DR PRINTS: PR00170; NACHANNEL.
 SQ SEQUENCE 2146 AA; 244649 MW; 281BECCAC494AE49 CRC64;

Query Match 100.0%; Score 34; DB 13; Length 2146;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 RHRGS 6
 DB 947 RHRGS 952

RESULT 10
 O9PM4 PRELIMINARY; PRT; 2157 AA.
 ID O9PM4;
 AC O9PM4;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE N-TYPE CALCIUM CHANNEL ALPHA-1B CDB8 VARIANT.
 GN CACNA1B.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID-9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-DORSAL ROOT GANGLION NEURONS;
 RX Lu Q., Dunlap K.;
 RT "Cloning and Functional Expression of Novel N-type Ca2+ Channel
 RT Variants.";
 RL J. Biol. Chem. 274:34566-34575(1999).
 DR EMBL: AF173019; AAD51822.1;
 DR InterPro: IPR000636; Cation_chan_non_119.
 DR InterPro: IPR001682; Channel_pore_Ca_Na.
 DR InterPro: IPR001696; Na_channel.
 DR InterPro: IPR002077; Ca_channel.
 DR InterPro: IPR002111; Cat_channel_TrpL.
 DR Pfam: PF00520; Ion_trans_4.
 DR PRINTS: PR00167; CACHANNEL.
 DR PRINTS: PR00170; NACHANNEL.
 SQ SEQUENCE 2157 AA; 245853 MW; 5E85C3AFA9493589 CRC64;

Query Match 100.0%; Score 34; DB 13; Length 2157;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRGS 6
 DB 958 RHRGS 963

RESULT 11
 O9PM6 PRELIMINARY; PRT; 2171 AA.
 ID O9PM6;
 AC O9PM6;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE N-TYPE CALCIUM CHANNEL ALPHA-1B CDB6 VARIANT.
 GN CACNA1B.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID-9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-DORSAL ROOT GANGLION NEURONS;
 RX Lu Q., Dunlap K.;
 RT "Cloning and Functional Expression of Novel N-type Ca2+ Channel
 RT Variants.";
 RL J. Biol. Chem. 274:34566-34575(1999).
 DR EMBL: AF173017; AAD51820.1;
 DR InterPro: IPR000636; Cation_chan_non_119.
 DR InterPro: IPR001682; Channel_pore_Ca_Na.
 DR InterPro: IPR001696; Na_channel.
 DR InterPro: IPR002077; Ca_channel.
 DR InterPro: IPR002111; Cat_channel_TrpL.
 DR Pfam: PF00520; Ion_trans_4.
 DR PRINTS: PR00167; CACHANNEL.

DR PRINTS: PR00170; NACHANNEL.
SQ SEQUENCE 2171 AA; 247055 MW; 4023DD3D1AF7EC3 CRC64;

Query Match 100.0%; Score 34; DB 13; Length 2171;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRGS 6
DB 972 RHRGS 977

RESULT 12

O9PUM5 PRELIMINARY; PRT; 2182 AA.

AC O9PUM5: 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE N-TYPE CALCIUM CHANNEL ALPHA-1B CDB2 VARIANT.
GN CACNA1B.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-DORSAL ROOT GANGLION NEURONS;
RX MEDLINE=20044702; PubMed=10574919;
RA Lu Q., Dunlap K.;
RT "Cloning and Functional Expression of Novel N-type Ca2+ Channel Variants."
RT J. Biol. Chem. 274:34566-34575(1999).
DR EMBL: AF173013; AAD51816.1;
DR InterPro: IPR000636; Cation_chan_non_119.
DR InterPro: IPR001682; Channel_pore_Ca_Na.
DR InterPro: IPR001696; Na_channel.
DR InterPro: IPR002077; Ca_channel.
DR InterPro: IPR002111; Cat_channel_TripL.
DR Pfam: PF00520; Ion_trans_4.
DR PRINTS: PR00167; CACCHANNEL.
DR PRINTS: PR00170; NACHANNEL.
SQ SEQUENCE 2182 AA; 248259 MW; 1C34F9DF5991783E CRC64;

Query Match 100.0%; Score 34; DB 13; Length 2182;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRGS 6
DB 983 RHRGS 988

RESULT 13

O9PUM4 PRELIMINARY; PRT; 2321 AA.

AC O9PUM4: 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE N-TYPE CALCIUM CHANNEL ALPHA-1B CDB3 VARIANT.
GN CACNA1B.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-DORSAL ROOT GANGLION NEURONS;

RX MEDLINE=20044702; PubMed=10574919;
RA Lu Q., Dunlap K.;
RT "Cloning and Functional Expression of Novel N-type Ca2+ Channel Variants."
RT J. Biol. Chem. 274:34566-34575(1999).
DR EMBL: AF173014; AAD51817.1;
DR InterPro: IPR000636; Cation_chan_non_119.
DR InterPro: IPR001682; Channel_pore_Ca_Na.
DR InterPro: IPR001696; Na_channel.
DR InterPro: IPR002077; Ca_channel.
DR InterPro: IPR002111; Cat_channel_TripL.
DR Pfam: PF00520; Ion_trans_4.
DR PRINTS: PR00167; CACCHANNEL.
DR PRINTS: PR00170; NACHANNEL.
SQ SEQUENCE 2321 AA; 262972 MW; E6FB221909BFD20B CRC64;

Query Match 100.0%; Score 34; DB 13; Length 2321;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRGS 6
DB 947 RHRGS 952

RESULT 14

O9PTA4 PRELIMINARY; PRT; 2331 AA.

AC O9PTA4: 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE N-TYPE CALCIUM CHANNEL ALPHA1B SUBUNIT.
GN CACNA1B.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Cahill A.L., Huxley J.H., Fox A.P.;
RT "Co-expression of cloned alpha1B, beta2a, and alpha2/delta subunits produces non-inactivating calcium channels similar to those found in bovine chromaffin cells."
RT J. Neurosci. 0:0-0(2000).
DR EMBL: AF173882; AAE24229.1;
DR InterPro: IPR000636; Cation_chan_non_119.
DR InterPro: IPR001682; Channel_pore_Ca_Na.
DR InterPro: IPR001696; Na_channel.
DR InterPro: IPR002077; Ca_channel.
DR InterPro: IPR002111; Cat_channel_TripL.
DR Pfam: PF00520; Ion_trans_4.
DR PRINTS: PR00167; CACCHANNEL.
DR PRINTS: PR00170; NACHANNEL.
SQ SEQUENCE 2331 AA; 260560 MW; 6A7F35C35C45ED8 CRC64;

Query Match 100.0%; Score 34; DB 6; Length 2331;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHRGS 6
DB 962 RHRGS 967

RESULT 15

O9PM45 PRELIMINARY; PRT; 2332 AA.

AC O9PM45: 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE N-TYPE CALCIUM CHANNEL ALPHA-1B CDB7 VARIANT.
 GN CACHA1B.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID:9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-DORSAL ROOT GANGLION NEURONS;
 RA Lu Q., Dunlap K.;
 RT "Cloning and Functional Expression of Novel N-type Ca2+ Channel
 RT Variants.";
 RL J. Biol. Chem. 0:0-0(1999).
 DR EMBL: AF173018; AAD51821.1;
 DR InterPro: IPR000636; Cation_chan_non_1lg.
 DR InterPro: IPR001682; Channel_pore_Ca_Na.
 DR InterPro: IPR001696; Na_channel.
 DR InterPro: IPR002077; Ca_channel.
 DR InterPro: IPR002111; Ca_channel_TrtL.
 DR Pfam: PF00520; Ion_trans_4.
 DR PRINTS: PRO0167; CACHANNEL.
 DR PRINTS: PRO0170; NACHANNEL.
 SQ SEQUENCE 2332 AA; 264176 MW; ECBI3B79544F0594 CRC64;

Query Match 100.0%; Score 34; DB 13; Length 2332;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 RHRGS 6
 Db 958 RHRGS 963

Search completed: March 18, 2002, 09:51:54
 Job time: 283 sec

DR WPI: 1997-526399/48.

XX PT Plasminogen activator-inhibitor type I mutant inhibits elastase - or
XX PT has high affinity for vitronectin, for therapeutic inhibition of
XX PT elastase or vitronectin-mediated cell attachment, migration etc.

PS Claim 4: Page -: 144pp; English.

XX CC This polypeptide comprises a preferred mutant, P4Aa (V343A), of
XX CC human plasminogen activator inhibitor (PAI-1) mature protein (see
XX CC also AAM31587). It has an Ala for Val substitution at amino acid
XX CC residue 343, i.e. in the reactive centre loop of the protein. The
XX CC mutant can be obtained by site-directed mutagenesis of PAI-1 DNA
XX CC (see AAM97303) and expression in a host cell. Claimed PAI-1 mutants
XX CC (see AAM26710-25) have at least one amino acid substitution at
XX CC position 343-350, especially at position 343 and/or 346, and may
XX CC have additional stabilising mutations comprising between one and
XX CC four of His-150, Thr-154, Leu-319 and Ile-354, and one or more of
XX CC Arg-333, Arg-335, Gly-331, Ile-372 and Leu-91. The mutants are
XX CC characterised by being resistant to inactivation by elastase,
XX CC plasminogen activator, plasmin, cathepsin G, chymase, gelatinases A
XX CC and B, stromelysin and collagenase, and by having a high affinity
XX CC for vitronectin (Vn) and are used, in vivo or in vitro, to inhibit
XX CC pathological distress syndrome, acute lung inflammation, congenital
XX CC alpha 1-antitrypsin deficiency, cystic fibrosis, atopic dermatitis,
XX CC pancreatitis, periodontal disease, arthritis and HIV infection).
XX CC They also inhibit Vn-stimulated cell attachment, migration and/or
XX CC migration-induced proliferation (specifically in atherosclerosis,
XX CC post-angioplasty restenosis, neointima formation after vascular
XX CC trauma, vascular graft restenosis, fibrosis (associated with
XX CC chronic inflammation or chemotherapy, or of the lung), wound
XX CC healing with scarring and fibrosis, deep vein thrombosis and any
XX CC condition involving pathological angiogenesis (all claimed). Since
XX CC the mutants are resistant to cleavage by the proteases to which
XX CC they bind, they are more effective therapeutic agents. They also
XX CC promote uptake and clearance of elastase (or its complex with PAI-1)
XX CC by receptors.
XX CC (NB, the amino acid sequence of the mutant protein was obtained by
XX CC adaptation of the wild-type sequence provided in Fig 4A of the
XX CC specification).

SO Sequence 379 AA;

Query Match 100.0%; Score 55; DB 18; Length 379;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RMAPEIIMDR 11
DB 346 rmapeilmdr 356

RESULT 2
ID AAM26715 standard; Protein; 379 AA.
XX AAM26715;
XX AAM26715;
XX 14-APR-1998 (first entry)
XX Plasminogen activator inhibitor mutant P4ASP (V343D).
XX DE Plasminogen activator inhibitor type 1; PAI-1; human;
XX KW elastase; inhibitor; vitronectin; cell attachment; cell migration;
XX KW cell proliferation; emphysema; adult respiratory distress syndrome;
XX KW acute lung inflammation; alpha 1-antitrypsin deficiency;
XX KW cystic fibrosis; atopic dermatitis; pancreatitis;
XX KW periodontal disease; arthritis; HIV; atherosclerosis; restenosis;
XX KW neointima; fibrosis; wound healing; tumour; metastasis; psoriasis;
XX KW thrombosis; angiogenesis; therapy.

XX OS Homo sapiens.
XX OS Synthetic.
XX PN W09739028-A1.
XX 23-OCT-1997.
XX 11-APR-1997; 97WO-US06071.
XX 12-APR-1996; 96US-0015299.
XX (AMNA-) AMERICAN NAT RED CROSS.
XX Lawrence DA, Stefansson SP;
XX WPI: 1997-526399/48.

XX PT Plasminogen activator-inhibitor type I mutant inhibits elastase - or
XX PT has high affinity for vitronectin, for therapeutic inhibition of
XX PT elastase or vitronectin-mediated cell attachment, migration etc.

PS Claim 4: Page -: 144pp; English.

XX CC This polypeptide comprises a preferred mutant, P4ASP (V343D), of
XX CC human plasminogen activator inhibitor (PAI-1) mature protein (see
XX CC also AAM31587). It has an Asp for Val substitution at amino acid
XX CC residue 343, i.e. in the reactive centre loop of the protein. The
XX CC mutant can be obtained by site-directed mutagenesis of PAI-1 DNA
XX CC (see AAM97303) and expression in a host cell. Claimed PAI-1 mutants
XX CC (see AAM26710-25) have at least one amino acid substitution at
XX CC position 343-350, especially at position 343 and/or 346, and may
XX CC have additional stabilising mutations comprising between one and
XX CC four of His-150, Thr-154, Leu-319 and Ile-354, and one or more of
XX CC Arg-333, Arg-335, Gly-331, Ile-372 and Leu-91. The mutants are
XX CC characterised by being resistant to inactivation by elastase,
XX CC plasminogen activator, plasmin, cathepsin G, chymase, gelatinases A
XX CC and B, stromelysin and collagenase, and by having a high affinity
XX CC for vitronectin (Vn) and are used, in vivo or in vitro, to inhibit
XX CC pathological distress syndrome, acute lung inflammation, congenital
XX CC alpha 1-antitrypsin deficiency, cystic fibrosis, atopic dermatitis,
XX CC pancreatitis, periodontal disease, arthritis and HIV infection).
XX CC They also inhibit Vn-stimulated cell attachment, migration and/or
XX CC migration-induced proliferation (specifically in atherosclerosis,
XX CC post-angioplasty restenosis, neointima formation after vascular
XX CC trauma, vascular graft restenosis, fibrosis (associated with
XX CC chronic inflammation or chemotherapy, or of the lung), wound
XX CC healing with scarring and fibrosis, deep vein thrombosis and any
XX CC condition involving pathological angiogenesis (all claimed). Since
XX CC the mutants are resistant to cleavage by the proteases to which
XX CC they bind, they are more effective therapeutic agents. They also
XX CC promote uptake and clearance of elastase (or its complex with PAI-1)
XX CC by receptors.
XX CC (NB, the amino acid sequence of the mutant protein was obtained by
XX CC adaptation of the wild-type sequence provided in Fig 4A of the
XX CC specification).

SO Sequence 379 AA;

Query Match 100.0%; Score 55; DB 18; Length 379;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RMAPEIIMDR 11
DB 346 rmapeilmdr 356

RESULT 3
ID AAM26716

ID AAM26716 standard; Protein; 379 AA.
 AC AAM26716;
 XX
 XX
 DT 14-APR-1998 (first entry)
 DE Plasminogen activator inhibitor mutant P4Gly (V343G).
 XX
 XX Plasminogen activator inhibitor type 1; PAI-1; human;
 KM elastase; inhibitor; vitronectin; cell attachment; cell migration;
 KM cell proliferation; emphysema; adult respiratory distress syndrome;
 KM acute lung inflammation; alpha 1-antitrypsin deficiency;
 KM cystic fibrosis; atopic dermatitis; pancreatitis;
 KM periodontal disease; arthritis; HIV; atherosclerosis; restenosis;
 KM neointima; fibrosis; wound healing; tumour; metastasis; psoriasis;
 KM thrombosis; angiogenesis; therapy.
 KM
 OS Homo sapiens.
 OS Synthetic.
 PN W09739028-A1.
 XX
 XX 23-OCT-1997.
 PD
 XX 11-APR-1997; 97WO-US06071.
 PF
 XX 12-APR-1996; 96US-0015299.
 PR
 XX (AMNA-) AMERICAN NAT RED CROSS.
 PA
 XX Lawrence DA, Stefansson SP;
 PI
 XX WPI: 1997-526399/48.
 DR
 XX Plasminogen activator-inhibitor type I mutant inhibits elastase - or
 PT has high affinity for vitronectin, for therapeutic inhibition of
 PT elastase or vitronectin-mediated cell attachment, migration etc.
 PS
 XX Claim 4; Page -: 144pp; English.

CC specification).
 XX
 XX Sequence 379 AA;
 SQ
 Query Match 100.0%; Score 55; DB 18; Length 379;
 Best Local Similarity 100.0%; Pred. No. 0.022;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RMAPEEIMDR 11
 DB 346 rmapeeimdr 356
 RESULT 4
 AAM26717
 ID AAM26717 standard; Protein; 379 AA.
 XX
 XX AAM26717;
 AC
 XX 14-APR-1998 (first entry)
 DT
 DE Plasminogen activator inhibitor mutant P4Leu (V343L).
 XX
 XX Plasminogen activator inhibitor type 1; PAI-1; human;
 KM elastase; inhibitor; vitronectin; cell attachment; cell migration;
 KM cell proliferation; emphysema; adult respiratory distress syndrome;
 KM acute lung inflammation; alpha 1-antitrypsin deficiency;
 KM cystic fibrosis; atopic dermatitis; pancreatitis;
 KM periodontal disease; arthritis; HIV; atherosclerosis; restenosis;
 KM neointima; fibrosis; wound healing; tumour; metastasis; psoriasis;
 KM thrombosis; angiogenesis; therapy.
 KM
 OS Homo sapiens.
 OS Synthetic.
 PN W09739028-A1.
 XX
 XX 23-OCT-1997.
 PD
 XX 11-APR-1997; 97WO-US06071.
 PF
 XX 12-APR-1996; 96US-0015299.
 PR
 XX (AMNA-) AMERICAN NAT RED CROSS.
 PA
 XX Lawrence DA, Stefansson SP;
 PI
 XX WPI: 1997-526399/48.
 DR
 XX Plasminogen activator-inhibitor type I mutant inhibits elastase - or
 PT has high affinity for vitronectin, for therapeutic inhibition of
 PT elastase or vitronectin-mediated cell attachment, migration etc.
 PS
 XX Claim 4; Page -: 144pp; English.

This polypeptide comprises a preferred mutant, P4Leu (V343L), of human plasminogen activator inhibitor (PAI-1) mature protein (see also AAM31587). It has a Leu for Val substitution at amino acid residue 343, i.e. in the reactive centre loop of the protein. The mutant can be obtained by site-directed mutagenesis of PAI-1 DNA (see AAT97303) and expression in a host cell. Claimed PAI-1 mutants (see AAM26710-25) have at least one amino acid substitution at position 343-350, especially at position 343 and/or 346, and may have additional stabilising mutations comprising between one and four of His-150, Thr-154, Leu-319 and Ile-354, and one or more of Arg-333, Arg-335, Gly-331, Ile-372 and Leu-91. The mutants are characterised by being resistant to inactivation by elastase, plasminogen activator, plasmin, cathepsin G, chymase, gelatinases A and B, stromelysin and collagenase, and by having a high affinity for vitronectin (Vn) and are used, in vivo or in vitro, to inhibit pathological elastase activity (specifically in emphysema, adult respiratory distress syndrome, acute lung inflammation, congenital alpha 1-antitrypsin deficiency, cystic fibrosis, atopic dermatitis, pancreatitis, periodontal disease, arthritis and HIV infection). They also inhibit Vn-stimulated cell attachment, migration and/or migration-induced proliferation (specifically in atherosclerosis, post-angioplasty restenosis, neointima formation after vascular trauma, vascular graft restenosis, fibrosis (associated with chronic inflammation or chemotherapy, or of the lung), wound healing with scarring and fibrosis, growth/invansion of primary tumours or metastases, psoriasis, deep vein thrombosis and any condition involving pathological angiogenesis (all claimed). Since the mutants are resistant to cleavage by the proteases to which they bind, they are more effective therapeutic agents. They also promote uptake and clearance of elastase (or its complex with PAI-1) by receptors.

(NB: the amino acid sequence of the mutant protein was obtained by adaptation of the wild-type sequence provided in Fig 4a of the

CC alpha 1-antitrypsin deficiency, cystic fibrosis, atopic dermatitis,
 CC pancreatitis, periodontal disease, arthritis and HIV infection).
 CC They also inhibit Vn-stimulated cell attachment, migration and/or
 CC migration-induced proliferation (specifically in atherosclerosis,
 CC post-angioplasty restenosis, neointima formation after vascular
 CC trauma, vascular graft restenosis, fibrosis (associated with
 CC chronic inflammation or chemotherapy, or of the lung), wound
 CC healing with scarring and fibrosis, growth/invasion of primary
 CC tumours or metastases, psoriasis, deep vein thrombosis and any
 CC condition involving pathological angiogenesis (all claimed). Since
 CC the mutants are resistant to cleavage by the proteases to which
 CC they bind, they are more effective therapeutic agents. They also
 CC promote uptake and clearance of elastase (or its complex with PAI-1)
 CC by receptors.
 CC (NB. the amino acid sequence of the mutant protein was obtained by
 CC adaptation of the wild-type sequence provided in Fig 4A of the
 CC specification).

SO Sequence 379 AA;

Query Match

Best Local Similarity 100.0%; Score 55; DB 18; Length 379;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RMAPEEIMDR 11
 DB 346 rmapeelindr 356

RESULT 5

AAW26718 standard; Protein; 379 AA.

XX AAW26718;

DT 14-APR-1998 (first entry)

DE Plasminogen activator inhibitor mutant P41le (V2431).

KW Plasminogen activator inhibitor type 1; PAI-1; human;
 KW elastase; inhibitor; vitronectin; cell attachment; cell migration;
 KW cell proliferation; emphysema; adult respiratory distress syndrome;
 KW acute lung inflammation; alpha 1-antitrypsin deficiency;
 KW cystic fibrosis; atopic dermatitis; pancreatitis;
 KW periodontal disease; arthritis; HIV; atherosclerosis; restenosis;
 KW neointima; fibrosis; wound healing; tumour; metastasis; psoriasis;
 KW thrombosis; angiogenesis; therapy.

OS Homo sapiens.
 OS Synthetic.

PM W09739028-A1.

PD 23-OCT-1997.

PF 11-APR-1997; 97WO-US06071.

PR 12-APR-1996; 96US-0015299.

PA (AMNA-) AMERICAN NAT RED CROSS.

PI Lawrence DA, Stefansson SP;

DR WPI; 1997-526399/48.

PT Plasminogen activator-inhibitor type I mutant inhibits elastase or
 PT has high affinity for vitronectin, for therapeutic inhibition of
 PT elastase or vitronectin-mediated cell attachment, migration etc.

PS Claim 4; Page -: 144p; English.

CC This polypeptide comprises a preferred mutant, P41le (V2431), of

CC human plasminogen activator inhibitor (PAI-1) mature protein (see
 CC also AAW31587). It has an ile for val substitution at amino acid
 CC residue 343, i.e. in the reactive centre loop of the protein. The
 CC mutant can be obtained by site-directed mutagenesis of PAI-1 DNA
 CC (see AAW37303) and expression in a host cell. Claimed PAI-1 mutants
 CC (see AAW6710-25) have at least one amino acid substitution at
 CC position 343-350 especially at position 343 and/or 346, and may
 CC have additional stabilising mutations comprising between one and
 CC four of His-150, Thr-154, Leu-319 and Ile-354, and one or more of
 CC Arg-333, Arg-335, Gly-331, Ile-372 and Leu-91. The mutants are
 CC characterised by being resistant to inactivation by elastase.
 CC Plasminogen activator, plasmin, cathepsin G, chymase, gelatinases A
 CC and B, stromelysin and collagenase, and by having a high affinity
 CC for vitronectin (Vn) and are used, in vivo or in vitro, to inhibit
 CC pathological distress activity (specifically in emphysema, adult
 CC respiratory distress syndrome, acute lung inflammation, congenital
 CC alpha 1-antitrypsin deficiency, cystic fibrosis, atopic dermatitis,
 CC pancreatitis, periodontal disease, arthritis and HIV infection).
 CC They also inhibit Vn-stimulated cell attachment, migration and/or
 CC migration-induced proliferation (specifically in atherosclerosis,
 CC post-angioplasty restenosis, neointima formation after vascular
 CC trauma, vascular graft restenosis, fibrosis (associated with
 CC chronic inflammation or chemotherapy, or of the lung), wound
 CC healing with scarring and fibrosis, deep vein thrombosis and any
 CC tumours or metastases, psoriasis, growth/invasion of primary
 CC condition involving pathological angiogenesis (all claimed). Since
 CC the mutants are resistant to cleavage by the proteases to which
 CC they bind, they are more effective therapeutic agents. They also
 CC promote uptake and clearance of elastase (or its complex with PAI-1)
 CC by receptors.
 CC (NB. the amino acid sequence of the mutant protein was obtained by
 CC adaptation of the wild-type sequence provided in Fig 4A of the
 CC specification).

SO Sequence 379 AA;

Query Match

Best Local Similarity 100.0%; Score 55; DB 18; Length 379;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RMAPEEIMDR 11
 DB 346 rmapeelindr 356

RESULT 6

AAW04922 standard; Protein; 379 AA.

XX AAW04922;

DT 26-SEP-2001 (first entry)

DE Human Plasminogen activator inhibitor-1, PAI-1, mutant P41A.

KW Human plasminogen activator inhibitor-1; PAI-1; serpin; P41A;
 KW immobilised enzyme; cystic fibrosis; acute respiratory distress syndrome;
 KW ARDS; HIV infection; Human immunodeficiency virus; prostate cancer;
 KW TNF-mediated inflammation; benign prostatic hypertrophy; mutant; mutein;
 KW V343A.

OS Synthetic.

OS Homo sapiens.

PH Key Location/Qualifiers

FT Region 332..351 /label-Reactive_centre_loop

PN W0200138560-A2.

PD 31-MAY-2001.

PF 22-NOV-2000; 2000MO-US32315.
 XX 22-NOV-1999; 99US-0167553.
 XX (AMNA-) AMERICAN NAT RED CROSS.
 PA Lawrence DA, Day D;
 XX WPI: 2001-441438/47.
 DR
 XX
 PT Detecting a functionally active form of an enzyme in a biological
 PT sample comprises contacting an enzyme inhibitor immobilised on a solid
 PT substrate -
 XX
 PS Disclosure; Page - : 69pp; English.

CC The sequence represents human plasminogen activator inhibitor-1.
 CC PAI-1, mutant P4Aa (V343A). PAI-1 is a serine proteinase inhibitor
 CC or serpin. The protein is used to demonstrate the method of the invention
 CC which comprises detecting a functionally active form of an enzyme in a
 CC biological sample by contacting an enzyme inhibitor immobilised on a
 CC solid substrate with the biological sample and measuring the binding of
 CC the enzyme inhibitor to the active form of the enzyme by a detectable
 CC label, where the enzyme inhibitor specifically forms a covalent bond or
 CC binds with a dissociation constant of 1×10^{-9} M or less with the active
 CC form of the enzyme. The present invention provides a sensitive method for
 CC the detection of a functionally active form of an enzyme in a biological
 CC sample. Human PAI-1 can be used to detect a number of enzymes including
 CC tissue plasminogen activator, urokinase, thrombin, plasmin, neutrophil
 CC elastase, pancreatic elastase, trypsin, chymotrypsin, cathepsin G and
 CC prostate specific antigen and as such can be used in methods to diagnose
 CC diseases such as cystic fibrosis, acute respiratory distress syndrome
 CC (ARDS), HIV infection, TNF-mediated inflammation, prostate cancer and
 CC benign prostatic hypertrophy.
 CC Note: The present sequence is not shown in the specification but is
 CC derived from the mature PAI-1 sequence shown in figure 5.
 CC
 XX
 SO Sequence 379 AA:

Query Match 100.0%; Score 55; DB 22; Length 379;
 Best Local Similarity 100.0%; Pred. No. 0.022;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RMAPEETIMDR 11
 |||||
 DB 346 rmapeelindr 356

RESULT 7
 AA004923
 ID AA004923 standard; Protein: 379 AA.
 AC AA004923:
 XX
 DT 26-SEP-2001 (first entry)
 XX
 DE Human Plasminogen activator inhibitor-1, PAI-1, mutant P4ASP.
 XX
 KW Human; Plasminogen activator inhibitor-1; PAI-1; serpin; P4ASP;
 KW Immobilised enzyme; cystic fibrosis; acute respiratory distress syndrome;
 KW ARDS; HIV infection; Human immunodeficiency virus; prostate cancer;
 KW TNF-mediated inflammation; benign prostatic hypertrophy; mutant; mutein;
 KW V343D.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 XX
 FT Key Location/Qualifiers
 FT 332..351
 FT Region /label- Reactive_centre_loop
 XX
 PN MO200138560-A2.

XX 31-MAY-2001.
 PD
 XX 22-NOV-2000; 2000MO-US32315.
 XX 22-NOV-1999; 99US-0167553.
 XX (AMNA-) AMERICAN NAT RED CROSS.
 PA Lawrence DA, Day D;
 XX WPI: 2001-441438/47.
 DR
 XX
 PT Detecting a functionally active form of an enzyme in a biological
 PT sample comprises contacting an enzyme inhibitor immobilised on a solid
 PT substrate -
 XX
 PS Disclosure; Page - : 69pp; English.

CC The sequence represents human plasminogen activator inhibitor-1.
 CC PAI-1, mutant P4ASP (V343D). PAI-1 is a serine proteinase inhibitor
 CC or serpin. The protein is used to demonstrate the method of the invention
 CC which comprises detecting a functionally active form of an enzyme in a
 CC biological sample by contacting an enzyme inhibitor immobilised on a
 CC solid substrate with the biological sample and measuring the binding of
 CC the enzyme inhibitor to the active form of the enzyme by a detectable
 CC label, where the enzyme inhibitor specifically forms a covalent bond or
 CC binds with a dissociation constant of 1×10^{-9} M or less with the active
 CC form of the enzyme. The present invention provides a sensitive method for
 CC the detection of a functionally active form of an enzyme in a biological
 CC sample. Human PAI-1 can be used to detect a number of enzymes including
 CC tissue plasminogen activator, urokinase, thrombin, plasmin, neutrophil
 CC elastase, pancreatic elastase, trypsin, chymotrypsin, cathepsin G and
 CC prostate specific antigen and as such can be used in methods to diagnose
 CC diseases such as cystic fibrosis, acute respiratory distress syndrome
 CC (ARDS), HIV infection, TNF-mediated inflammation, prostate cancer and
 CC benign prostatic hypertrophy.
 CC Note: The present sequence is not shown in the specification but is
 CC derived from the mature PAI-1 sequence shown in figure 5.
 CC
 XX
 SO Sequence 379 AA:

Query Match 100.0%; Score 55; DB 22; Length 379;
 Best Local Similarity 100.0%; Pred. No. 0.022;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RMAPEETIMDR 11
 |||||
 DB 346 rmapeelindr 356

RESULT 8
 AA004924
 ID AA004924 standard; Protein: 379 AA.
 AC AA004924:
 XX
 DT 26-SEP-2001 (first entry)
 XX
 DE Human Plasminogen activator inhibitor-1, PAI-1, mutant P4Gly.
 XX
 KW Human; Plasminogen activator inhibitor-1; PAI-1; serpin; P4Gly;
 KW Immobilised enzyme; cystic fibrosis; acute respiratory distress syndrome;
 KW ARDS; HIV infection; Human immunodeficiency virus; prostate cancer;
 KW TNF-mediated inflammation; benign prostatic hypertrophy; mutant; mutein;
 KW V343G.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 XX
 FT Key Location/Qualifiers
 FT 332..351
 FT Region


```

FT      /label= Reactive_centre_loop
PM      WO200138560-A2.
XX
PD      31-MAY-2001.
XX
XX      22-NOV-2000; 2000WO-US32315.
XX
PR      22-NOV-1999; 99US-0167553.
PA      (AMNA-) AMERICAN NAT RED CROSS.
PI      Lawrence DA, Day D:
DR      WPI: 2001-441438/47.
XX
XX      Detecting a functionally active form of an enzyme in a biological
PT      sample comprises contacting an enzyme inhibitor immobilised on a solid
PT      substrate.
XX
XX      Disclosure: Page - : 69pp: English.
XX
XX      The sequence represents human plasminogen activator inhibitor-1,
CC      PAI-1, mutant P40ly (V343G). PAI-1 is a serine proteinase inhibitor
CC      or serpin. The protein is used to demonstrate the method of the invention
CC      which comprises detecting a functionally active form of an enzyme in a
CC      biological sample by contacting an enzyme inhibitor immobilised on a
CC      solid substrate with the biological sample and measuring the binding of
CC      the enzyme inhibitor to the active form of the enzyme by a detectable
CC      label, where the enzyme inhibitor specifically forms a covalent bond or
CC      binds with a dissociation constant of 1 x 10-9M or less with the active
CC      form of the enzyme. The present invention provides a sensitive method for
CC      the detection of a functionally active form of an enzyme in a biological
CC      sample. Human PAI-1 can be used to detect a number of enzymes including
CC      tissue plasminogen activator, urokinase, thrombin, plasmin, neutrophil
CC      elastase, pancreatic elastase, trypsin, chymotrypsin, cathepsin G and
CC      prostate specific antigen and as such can be used in methods to diagnose
CC      diseases such as cystic fibrosis, acute respiratory distress syndrome
CC      (ARDS), HIV infection, TNF-mediated inflammation, prostate cancer and
CC      benign prostatic hypertrophy.
CC      Note: The present sequence is not shown in the specification but is
CC      derived from the mature PAI-1 sequence shown in figure 5.
XX
XX      Sequence 379 AA:
XX
XX      Query Match 100.0%; Score 55; DB 22; Length 379;
XX      Best Local Similarity 100.0%; Pred. No. 0.022; 0; Indels 0; Gaps 0;
XX      Matches 11; conservative 0; mismatches 0;
XX
XX      1 RMAPETIMDR 11
XX      |||||||||
XX      Db 346 rmapeelindr 356
XX
XX      RESULT 9
XX      AAU04925
XX      AAU04925 standard; Protein; 379 AA.
XX
XX      AAU04925;
XX
XX      26-SEP-2001 (first entry)
XX
XX      Human Plasminogen activator inhibitor-1, PAI-1, mutant P40eu.
XX
XX      Human: Plasminogen activator inhibitor-1; PAI-1; serpin; PA1eu:
XX      immobilised enzyme; cystic fibrosis; acute respiratory distress syndrome;
XX      AIDS; HIV infection; Human immunodeficiency Virus; prostate cancer;
XX      TNF-mediated inflammation; benign prostatic hypertrophy; mutant;
XX      V343L.
XX
XX      Synthetic.
XX      Homo sapiens.
XX

```

FH	Key	Location/Qualifiers
FT	Region	332..351
XX	XX	/label- Reactive_centre_loop
XX	XX	
PN	MO200138560-A2.	
XX	XX	
PD	31-MAY-2001.	
PF	22-NOV-2000; 2000WO-US32315.	
PR	22-NOV-1999; 99US-0167553.	
PA	(AMNA-) AMERICAN NAT RED CROSS.	
PI	Lawrence DA, Day D;	
DR	WPI; 2001-441438/47.	
XX	XX	
PT	Detecting a functionally active form of an enzyme in a biological sample comprises contacting an enzyme inhibitor immobilised on a solid substrate -	
FT	disclosure; Page - ; 69pp; English.	
PS	XX	
XX	XX	
CC	The sequence represents human plasminogen activator inhibitor-1, PAI-1, mutant P4luu (V343L). PAI-1 is a serine proteinase inhibitor or serpin. Human PAI-1 is used to demonstrate the method of the invention which comprises detecting a functionally active form of an enzyme in a biological sample by contacting an enzyme inhibitor immobilised on a solid substrate with the biological sample and measuring the binding of the enzyme inhibitor to the active form of the enzyme by a detectable label, where the enzyme inhibitor specifically forms a covalent bond or binds with a dissociation constant of 1 x 10 ⁻⁹ M or less with the active form of the enzyme. The present invention provides a sensitive method for the detection of a functionally active form of an enzyme in a biological sample. Human PAI-1 can be used to detect a number of enzymes including tissue plasminogen activator, urokinase, thrombin, plasmin, neutrophil elastase, pancreatic elastase, trypsin, chymotrypsin, cathepsin G and prostate specific antigen and as such can be used in methods to diagnose diseases such as cystic fibrosis, acute respiratory distress syndrome (ARDS), HIV infection, TNF-mediated inflammation, prostate cancer and benign prostatic hypertrophy.	
CC	Note: The present sequence is not shown in the specification but is derived from the mature PAI-1 sequence shown in figure 5.	
CC	XX	
SQ	Sequence 379 AA:	
	Query Match 100.0%; Score 55; DB 22; Length 379; Best Local Similarity 100.0%; Pred. NO. 0.022; Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0 QY 1 RMAPEEIIIMDR 11 Db 346 rmapeeIImdr 356	
RESULT 10		
AU04926		
ID AU04926 standard; Protein; 379 AA.		
AC XX	AAU04926;	
DT XX	26-SEP-2001 (first entry)	
DE	Human Plasminogen activator inhibitor-1, PAI-1, mutant P4luu.	
KM	Human; Plasminogen activator inhibitor-1; PAI-1; serpin; P4luu; Immobilised enzyme; cystic fibrosis; acute respiratory distress syndrome; ARDS; HIV infection; Human immunodeficiency virus; prostate cancer; TNF-mediated inflammation; benign prostatic hypertrophy; mutant; muteln; V343L.	

```

XX OS Synthetic.
OS OS Homo sapiens.
XX XX
XX FH Key Location/Qualifiers
XX FT 332..351
XX FT Region /label= Reactive_centre_loop
XX PN WO200138560-A2.
XX PD 31-MAY-2001.
XX PF 22-NOV-2000; 2000MO-US32315.
XX PR 22-NOV-1999; 99US-0167553.
XX PA (AMNA-) AMERICAN NAT RED CROSS.
XX PI Lawrence DA, Day D;
XX DR WPI; 2001-441438/47.
XX XX
XX PT Detecting a functionally active form of an enzyme in a biological
XX PT sample comprises contacting an enzyme inhibitor immobilised on a solid
XX PT substrate -
XX PS Disclosure; Page - : 69pp; English.
XX CC The sequence represents human plasminogen activator inhibitor-1.
XX CC PAI-1, mutant PA11 (V343I). PAI-1 is a serine proteinase inhibitor
XX CC or serpin. The protein is used to demonstrate the method of the invention
XX CC which comprises detecting a functionally active form of an enzyme in a
XX CC biological sample by contacting an enzyme inhibitor immobilised on a
XX CC solid substrate with the biological sample and measuring the binding of
XX CC the enzyme inhibitor to the active form of the enzyme by a detectable
XX CC label, where the enzyme inhibitor specifically forms a covalent bond or
XX CC binds with a dissociation constant of 1 x 10-9M or less with the active
XX CC form of the enzyme. The present invention provides a sensitive method for
XX CC the detection of a functionally active form of an enzyme in a biological
XX CC sample. Human PAI-1 can be used to detect a number of enzymes including
XX CC tissue plasminogen activator, urokinase, thrombin, plasmin, neutrophil
XX CC elastase, pancreatic elastase, trypsin, chymotrypsin, cathepsin G and
XX CC prostate specific antigen and as such can be used in methods to diagnose
XX CC diseases such as cystic fibrosis, acute respiratory distress syndrome
XX CC (ARDS), HIV infection, TNF-mediated inflammation, prostate cancer and
XX CC benign prostatic hypertrophy.
XX CC Note: The present sequence is not shown in the specification but is
XX CC derived from the mature PAI-1 sequence shown in figure 5.
XX SO Sequence 379 AA:

Query Match 100.0%; Score 55; DB 22; Length 379;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RMAPEITMDR 11
DB 346 rmapeitmdr 356

RESULT 11
AAR08411
ID AAR08411 standard; protein; 382 AA.
XX AC AAR08411;
XX DT 27-FEB-1991 (first entry)
XX DE Modified mature Plasminogen Activator Inhibitor PAI-1.
XX KM Plasminogen Activator Inhibitor; fibrinolysis; liver transplants.
XX XX

```

```

OS OS Homo sapiens.
XX XX
XX PN MO9013648-A.
XX PD 15-NOV-1990.
XX PF 08-MAY-1990; 90MO-US02452.
XX PR 11-MAY-1989; 89US-0350264.
XX PA (DUPO ) DU PONT DE NEMOURS CO.
XX PI Davis GL, Knabb RM, Reilly TM, Sisk WP;
XX DR WPI; 1990-361483/48.
XX XX
XX PT New recombinant functional human plasminogen activator inhibitor
XX PT - and plasmid expression vectors for its prodn. in E. coli, used
XX PT for assays of T-PA PAI-1 inhibitors and treatment of excessive
XX PT fibrinolysis
XX PS Claim 30; Page 43; 49pp; English.
XX XX
XX CC The N-terminal sequence of this protein can be deleted to leave
XX CC SVH or VH. The Val residue of mature PAI-1 corresponds to residue
XX CC 24 of the preproteins. It may alternatively begin with. The protein
XX CC is used in tests to identify inhibitors of its binding to tissue
XX CC plasminogen activator. It is also used therapeutically to treat
XX CC excessive or inappropriate fibrinolysis, e.g. during liver
XX CC transplant surgery.
XX CC See also AA006590-006594.
XX SO Sequence 382 AA:

Query Match 100.0%; Score 55; DB 11; Length 382;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RMAPEITMDR 11
DB 349 rmapeitmdr 359

RESULT 12
AAP82007
ID AAP82007 standard; protein; 383 AA.
XX AC AAP82007;
XX DT 16-OCT-1990 (first entry)
XX DE Beta plasminogen activator inhibitor encoded by clone lambda 3.
XX KM Endothelial plasminogen activator inhibitor.
XX OS Homo sapiens.
XX XX
XX FH Key Location/Qualifiers
XX FT 1..4
XX FT label-signal sequence
XX FT 5..383
XX FT label-mat_beta-PAI
XX FT 209..211
XX FT Modified-site label-putative glycosylation site
XX FT 265..267
XX FT label-putative glycosylation site
XX FT 329..331
XX FT label-putative glycosylation site
XX PN MO8801273-A.
XX PD 25-FEB-1988.

```

XX 19-AUG-1987; 87WO-US02053.
 XX 19-AUG-1986; 86US-0897990.
 XX (SCRI-) SCRIPPS CLINIC AND RESEARCH FOUNDATION.
 XX Loskutoff DJ, Ny T, Sawley M;
 XX WPI: 1988-063992/09.
 XX N-PSDB; AAN80253.
 XX Diagnostic assays for tissue - and urokinase-type plasminogen
 XX inhibitor - and pure, recombinant human endothelial plasminogen
 XX activator inhibitor.
 XX Claim 11: Fig 22: 100pp; English.
 XX The lambda 3 clone DNA or plasmid PAI3 can be used to produce the
 XX inhibitor by recombinant DNA technology. The protein inhibits both
 XX t-PA and u-PA and is immunologically different from protease nexin
 XX and placental PAI. It can also be used in a solid phase assay for
 XX detection of endothelial PAI, and as an immunogen to raise antibodies
 XX for use as receptor molecules in other assay systems.
 XX See also AAR0252-54.
 XX Sequence 383 AA:

Query Match 100.0%; Score 55; DB 9; Length 383;
 Best Local Similarity 100.0%; Pred. No. 0.023;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RMAPEEIMDR 11
 DB 350 rmapeelindr 360

RESULT 13
 ID AAR81179 standard; protein; 402 AA.
 XX AAR81179;
 XX 12-DEC-1990 (first entry)
 XX Sequence of human endothelial plasminogen activator inhibitor (PAI)
 XX encoded by PAI2350 cDNA.
 XX Protease inhibitor.
 XX Homo sapiens.
 XX EP260757-A.
 XX 23-MAR-1988.
 XX 11-SEP-1987; 87EP-0201745.
 XX 12-SEP-1986; 86NL-0002307.
 XX (STCL) STICHT INSTAND BLOEDTRAN.
 XX Pannekoek H;
 XX WPI: 1988-079089/12.
 XX N-PSDB; AAN81524.
 XX Recombinant DNA coding human endothelial plasminogen activator inhibitor
 XX used for obtaining prods. having effect on blood coagulation or
 XX fibrinolysis
 XX Disclosure; / P; English.

XX Colonies of human endothelial cDNA were screened using the probe
 XX (AAN81522) and a heterologous antiserum to detect the presence of human
 XX endothelial PAI like polypeptides. Novel recombinant DNA molecule for use
 XX in cloning end/or expressing DNA sequence in host cells which contains,
 XX in addition to a vector protein, a DNA sequence coding for human PAI 1s
 XX claimed (see AAN81524).
 XX Sequence 402 AA:

Query Match 100.0%; Score 55; DB 9; Length 402;
 Best Local Similarity 100.0%; Pred. No. 0.024;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RMAPEEIMDR 11
 DB 369 rmapeelindr 379

RESULT 14
 ID AAR07986 standard; protein; 402 AA.
 XX AAR07986;
 XX 27-FEB-1991 (first entry)
 XX Plasminogen Activator Inhibitor PAI-1 encoded by clone ECE3-1.
 XX Plasminogen Activator Inhibitor; fibrinolysis; liver transplants.
 XX Homo sapiens.

Key Location/Qualifiers
 FT Peptide 1..23
 FT /label- leader peptide
 FT 24..402
 FT /label- mature PAI-1
 FT /note- "Met residue present as residue #1 of mature
 PAI-1"

XX WO9013648-A.
 XX 15-NOV-1990.
 XX 08-MAY-1990; 90WO-US02452.
 XX 11-MAY-1989; 89US-0350264.
 XX (DUPO) DU PONT DE NEMOURS CO.
 XX Davis GL, Knabb RM, Rellily TM, Slisk WP;
 XX WPI: 1990-361483/48.
 XX N-PSDB; AAO06594.

XX New recombinant functional human plasminogen activator inhibitor
 XX - and plasmid expression vectors for its prodn. in E. coli, used
 XX for assays of t-PA PAI-1 inhibitors and treatment of excessive
 XX fibrinolysis
 XX Disclosure; Page 8; 49pp; English.
 XX A human umbilical cord endothelial cell cDNA library was screened
 XX with 4 oligonucleotide probes based on the known PAI-1 sequence. A
 XX 2.1kb insert was identified in the most strongly positive clones.
 XX It was inserted into pTZ198 to give pCE3-1 and the sequence given
 XX here. To produce mature protein, the leader sequence
 XX was removed by partial digestion with AapI and NsiI of HaeIII-
 XX methyase treated pCE3-1 DNA. The resulting 1800bp fragment was
 XX attached to an adaptor (to restore the first 2, Met-Val, amino
 XX acids), digested with NcoI and inserted onto NcoI-PstI cleaved

CC PKR233-2 to give plasmid pVAC-PAI. This expresses mature PAI-1 with
 CC the N-terminus Val-His-His which is useful in tests to identify
 CC inhibitors of its binding to tissue plasminogen activator. The
 CC protein is also used therapeutically to treat excessive or
 CC inappropriate fibrinolysis, e.g. during liver transplant surgery.
 CC See also AA06590-006593.

XX Sequence 402 AA;

Query Match 100.0%; Score 55; DB 11; Length 402;
 Best Local Similarity 100.0%; Pred. No. 0.024;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RMAPEITMDR 11
 |||||
 DB 369 rmapelindr 379

RESULT 15

AAW31587
 ID AAW31587 standard; Protein; 402 AA.

XX AAW31587;

DT 14-APR-1998 (first entry)

DE Human plasminogen activator inhibitor type 1.

XX Plasminogen activator inhibitor type 1; PAI-1; human;
 KW elastase inhibitor; vitronectin; cell attachment; cell migration;
 KW cell proliferation; emphysema; adult respiratory distress syndrome;
 KW acute lung inflammation; alpha 1-antitrypsin deficiency;
 KW cystic fibrosis; atopic dermatitis; pancreatitis;
 KW periodontal disease; arthritis; HIV; atherosclerosis; restenosis;
 KW neointima; fibrosis; wound healing; tumour; metastasis; psoriasis;
 KW thrombosis; angiogenesis; therapy.

XX Homo sapiens.

PH Key Location/Qualifiers

FT Peptide 1..23

FT Protein 24..402

FT /label- Mat-Protein

FT /note- "Claim 3"

FT Region 354..375

FT /note- "reactive centre loop region"

FT Misc-difference 173

FT /note- "preferred substitution site for protein stabilisation"

FT Misc-difference 177

FT /note- "preferred substitution site for protein stabilisation"

FT Misc-difference 342

FT /note- "preferred substitution site for protein stabilisation"

FT Misc-difference 366

FT /label- P4

FT /note- "preferred substitution site to provide elastase inhibitor mutant"

FT Misc-difference 369

FT /label- P1

FT /note- "preferred substitution site to provide elastase inhibitor mutant"

FT Misc-difference 377

FT /note- "preferred substitution site for protein substitution"

XX WO9739028-A1.

XX 23-OCT-1997.

PF 11-APR-1997; 97WO-US06071.

XX 12-APR-1996; 96US-0015299.

XX (AMNA-) AMERICAN NAT RED CROSS.

PI Lawrence DA, Stefansson SP;

DR WPI: 1997-526399/48.

DR N-PSDB: AAT97303.

XX Plasminogen activator-inhibitor type I mutant inhibits elastase - or
 PT has high affinity for vitronectin, for therapeutic inhibition of
 PT elastase or vitronectin-mediated cell attachment, migration etc.

PS Claim 3; Fig 4A; 144pp; English.

XX This polypeptide sequence comprises wild-type human plasminogen
 CC activator inhibitor type (PAI-1). Novel mutants (see AAW26710-25) of
 CC the PAI-1 mature protein are claimed that inhibit elastase or other
 CC elastase-like proteases, or are inhibitors of vitronectin-dependent
 CC cell migration. The mutants are obtained by site-directed
 CC mutagenesis of the PAI-1 DNA sequence (see AAT97303) and expression
 CC in host cells, and have a range of therapeutic uses. Preferred
 CC mutants have amino acid substitutions in the reactive centre
 CC loop region (especially at position 343 and/or 346 of the mature
 CC protein), and may have additional stabilising amino acid
 CC substitutions at 1-4 of residues 150, 154, 319 and 354, and 1-5
 CC of residues 333, 335, 331, 372 and 91.

XX Sequence 402 AA;

Query Match 100.0%; Score 55; DB 18; Length 402;
 Best Local Similarity 100.0%; Pred. No. 0.024;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RMAPEITMDR 11
 |||||

DB 369 rmapelindr 379

Search completed: March 18, 2002, 09:48:55
 Job time: 364 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:49:40 ; Search time 37.77 seconds
(without alignments)
6.554 Million cell updates/sec

Title: US-09-544-665-2

Sequence: 1 RMAPEEIMDR 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents, PA.*
1: /cgn2_6/pdata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/pdata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/pdata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/pdata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/pdata/2/1aa/PCFUS_COMB.pep.*
6: /cgn2_6/pdata/2/1aa/PCFUS1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	55	100.0	379 3 US-08-840-204-3	Sequence 3, Appl
2	55	100.0	390 1 US-08-121-714-6	Sequence 6, Appl
3	55	100.0	390 1 US-08-477-108A-6	Sequence 6, Appl
4	55	100.0	390 2 US-08-477-112-6	Sequence 6, Appl
5	55	100.0	390 5 PCF-US93-08322-6	Sequence 6, Appl
6	55	100.0	402 1 US-08-315-461-7	Sequence 7, Appl
7	55	100.0	402 3 US-08-948-097-4	Sequence 4, Appl
8	55	100.0	402 3 US-08-840-204-2	Sequence 2, Appl
9	55	100.0	402 4 US-09-348-817A-4	Sequence 4, Appl
10	55	100.0	402 4 US-09-026-408-3	Sequence 3, Appl
11	50	90.9	33 1 US-08-781-020-6	Sequence 6, Appl
12	50	90.9	33 4 US-09-038-935-6	Sequence 6, Appl
13	34	61.8	37 1 US-08-041-774-3	Sequence 2, Appl
14	34	61.8	37 4 US-08-530-340-2	Sequence 2, Appl
15	34	61.8	216 1 US-08-041-774-5	Sequence 5, Appl
16	34	61.8	216 4 US-08-530-340-5	Sequence 5, Appl
17	34	61.8	216 5 PCF-US96-12129B-2	Sequence 2, Appl
18	34	61.8	757 4 US-09-413-814-84	Sequence 84, Appl
19	33	60.0	471 4 US-08-961-083-158	Sequence 158, Appl
20	33	60.0	487 4 US-09-058-947A-4	Sequence 4, Appl
21	33	60.0	500 4 US-08-868-373-12	Sequence 12, Appl
22	33	60.0	753 3 US-08-674-887A-6	Sequence 6, Appl
23	33	60.0	753 3 US-08-951-844-6	Sequence 6, Appl
24	32	58.2	166 4 US-08-954-395A-10	Sequence 10, Appl
25	32	58.2	166 4 US-08-954-395A-15	Sequence 15, Appl
26	32	58.2	166 4 US-08-954-395A-16	Sequence 16, Appl
27	32	58.2	166 4 US-08-954-395A-17	Sequence 17, Appl

28	32	58.2	260 2 US-07-857-224B-7	Sequence 7, Appl
29	32	58.2	405 4 US-09-232-200-63	Sequence 63, Appl
30	32	58.2	405 4 US-09-232-197-63	Sequence 63, Appl
31	32	58.2	429 2 US-08-677-049-5	Sequence 5, Appl
32	32	58.2	468 2 US-08-959-011-3	Sequence 3, Appl
33	32	58.2	506 4 US-09-232-191-9	Sequence 9, Appl
34	32	58.2	506 4 US-09-232-200-9	Sequence 9, Appl
35	32	58.2	506 4 US-09-232-200-95	Sequence 95, Appl
36	32	58.2	506 4 US-09-232-197-9	Sequence 9, Appl
37	32	58.2	506 4 US-09-232-197-95	Sequence 95, Appl
38	32	58.2	632 4 US-09-232-200-35	Sequence 35, Appl
39	32	58.2	632 4 US-09-232-197-35	Sequence 35, Appl
40	32	58.2	643 4 US-09-232-200-42	Sequence 42, Appl
41	32	58.2	643 4 US-09-232-200-45	Sequence 45, Appl
42	32	58.2	643 4 US-09-232-200-71	Sequence 71, Appl
43	32	58.2	643 4 US-09-232-197-42	Sequence 42, Appl
44	32	58.2	643 4 US-09-232-197-45	Sequence 45, Appl
45	32	58.2	643 4 US-09-232-197-71	Sequence 71, Appl

ALIGNMENTS

```
RESULT 1
US-08-840-204-3
; Sequence 3, Application US/08840204
; Patent No. 6103498
;
GENERAL INFORMATION:
APPLICANT: LAWRENCE, DANIEL A.
APPLICANT: STEFANSSON, STEINGRIMUR P.
TITLE OF INVENTION: MUTANT PLASMINOGEN ACTIVATOR-INHIBITOR
TITLE OF INVENTION: TYPE I (PAT-1) AND USES THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESSES:
ADDRESS: MORRISON & FOERSTER
STREET: 2000 PENNSYLVANIA AVENUE, NW
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/840,204
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: LYNAT, SHMUEL
REGISTRATION NUMBER: 33,949
REFERENCE/DOCKET NUMBER: 30807-20004.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 822-0168
TELEX: 90-4030 MRSNFOERSM
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 379 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-840-204-3

Query Match 100.0%; Score 55; DB 3; Length 379;
Best Local Similarity 100.0%; Pred No. 0.008;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RMAPEEIMDR 11
Db 346 RMAPEEIMDR 356
```

RESULT 2
US-08-121-714-6
Sequence 6, Application US/08121714
Patent No. 5470970
GENERAL INFORMATION:
APPLICANT: Sager, Ruth
TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH
TITLE OF INVENTION: TUMOR SUPPRESSING ACTIVITY
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/121,714
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/938,823
FILING DATE: 09/01/92
APPLICATION NUMBER: 07/844,296
FILING DATE: 02/28/92
APPLICATION NUMBER: 07/662,216
FILING DATE: 02/28/91
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 00530/072001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 390
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-121-714-6

Query Match 100.0%; Score 55; DB 1; Length 390;
Best Local Similarity 100.0%; Pred. NO. 0.0083;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RMAPEIIMDR 11
|||||
Db 369 RMAPEIIMDR 379

RESULT 3
US-08-477-108A-6
Sequence 6, Application US/08477108A
Patent No. 5801001
GENERAL INFORMATION:
APPLICANT: Sager, Ruth
APPLICANT: Zou, Zhigang
APPLICANT: Anisowicz, Anthony
TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH
TITLE OF INVENTION: TUMOR SUPPRESSING ACTIVITY
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,108A
FILING DATE: June 7, 1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/121,714
FILING DATE: 09/01/93
APPLICATION NUMBER: 07/938,823
FILING DATE: 09/01/92
APPLICATION NUMBER: 07/844,296
FILING DATE: 02/28/92
APPLICATION NUMBER: 07/662,216
FILING DATE: 02/28/91
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 06570/002002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 390
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-477-108A-6

Query Match 100.0%; Score 55; DB 1; Length 390;
Best Local Similarity 100.0%; Pred. NO. 0.0083;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RMAPEIIMDR 11
|||||
Db 369 RMAPEIIMDR 379

RESULT 4
US-08-477-112-6
Sequence 6, Application US/08477112
Patent No. 5905023
GENERAL INFORMATION:
APPLICANT: Sager, Ruth
TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH
TITLE OF INVENTION: TUMOR SUPPRESSING ACTIVITY
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,112
FILING DATE:

CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/121,714
FILING DATE: 09/01/93
APPLICATION NUMBER: 07/938,823
FILING DATE: 09/01/92
APPLICATION NUMBER: 07/844,296
FILING DATE: 02/28/92
APPLICATION NUMBER: 07/662,216
FILING DATE: 02/28/91
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 06570/002003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 390
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-477-112-6

Query Match 100.0% Score 55; DB 2; Length 390;
Best Local Similarity 100.0%; Pred. No. 0.0083;
Matches 11: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RMAPEEIIIMDR 11
|||||
DB 369 RMAPEEIIIMDR 379

RESULT 5
PCT-US93-08322-6
Sequence 6, Application PC/TUS9308322
GENERAL INFORMATION:
APPLICANT: Sager, Ruth
TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH TUMOR SUPPRESSING ACTIVITY
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/08322
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/938,823
FILING DATE: 09/01/92
APPLICATION NUMBER: 07/844,296
FILING DATE: 02/28/92
APPLICATION NUMBER: 07/662,216
FILING DATE: 02/28/91
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 00530/072001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906

TELEX: 200154
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 390
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
PCT-US93-08322-6

Query Match 100.0% Score 55; DB 5; Length 390;
Best Local Similarity 100.0%; Pred. No. 0.0083;
Matches 11: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RMAPEEIIIMDR 11
|||||
DB 369 RMAPEEIIIMDR 379

RESULT 6
US-08-315-461-7
Sequence 7, Application US/08315461
Patent No. 5639726
GENERAL INFORMATION:
APPLICANT: Lawrence, Daniel A.
APPLICANT: Ginsburg, David
APPLICANT: Shore, Joseph D.
APPLICANT: Fay, William P.
APPLICANT: Olson, Steven T.
APPLICANT: Francis-Chmura, Anne-Marie
APPLICANT: Daniel T. Eltzman
APPLICANT: Dell Palelli
TITLE OF INVENTION: Peptide Mediated Enhancement of
TITLE OF INVENTION: Thrombolysis: Methods and Compositions
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/315,461
FILING DATE: Concurrently Herewith
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UMIC.006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 402 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-315-461-7

Query Match 100.0% Score 55; DB 1; Length 402;
Best Local Similarity 100.0%; Pred. No. 0.0083;
Matches 11: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RMAPEEIIIMDR 11

Db 369 RMAPEIIMDR 379

RESULT 7
US-08-948-997-4
Sequence 4, Application US/08948997

Patent No. 6008020
GENERAL INFORMATION:
APPLICANT: HASTINGS, GREGG
APPLICANT: COLEMAN, TIM
TITLE OF INVENTION: BRAIN-ASSOCIATED INHIBITOR OF
TITLE OF INVENTION: TISSUE-TYPE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: USA

ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/948,997
FILING DATE: Oct-10-97
CLASSIFICATION: 439

ATTORNEY/AGENT INFORMATION:
NAME: A. ANDERS BROOKES
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF336
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 402 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-948-997-4

Query Match 100.0%; Score 55; DB 3; Length 402;
Best Local Similarity 100.0%; Pred. No. 0.0085;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RMAPEIIMDR 11
Db 369 RMAPEIIMDR 379

RESULT 8
US-08-840-204-2
Sequence 2, Application US/08840204

Patent No. 6103498
GENERAL INFORMATION:
APPLICANT: LAWRENCE, DANIEL A.
APPLICANT: STEFANSON, STEINGRIMUR P.
TITLE OF INVENTION: MUTANT PLASMINOGEN ACTIVATOR-INHIBITOR
TITLE OF INVENTION: TYPE 1 (PAI-1) AND USES THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 PENNSYLVANIA AVENUE, NW
CITY: WASHINGTON
STATE: DC
COUNTRY: USA

ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/840,204
FILING DATE:
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: LIVNAT, SHMOEL
REGISTRATION NUMBER: 33,949
REFERENCE/DOCKET NUMBER: 30807-20004,00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 822-0168
TELEX: 90-4030 MRSNFOERSMWSH
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 402 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-840-204-2

Query Match 100.0%; Score 55; DB 3; Length 402;
Best Local Similarity 100.0%; Pred. No. 0.0085;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RMAPEIIMDR 11
Db 369 RMAPEIIMDR 379

RESULT 9
US-09-348-817A-4
Sequence 4, Application US/09348817A
Patent No. 6191260
GENERAL INFORMATION:
APPLICANT: Hastings et al.
TITLE OF INVENTION: Brain-Associated Inhibitor of Tissue-Type Plasminogen
TITLE OF INVENTION: Activator
FILE REFERENCE: PF336D1
CURRENT APPLICATION NUMBER: US/09/348,817A
CURRENT FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: 08/948,997
PRIOR FILING DATE: 1997-10-10
PRIOR APPLICATION NUMBER: 60/028,117
PRIOR FILING DATE: 1996-10-11
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 4
LENGTH: 402
TYPE: PRT
ORGANISM: Bos taurus
US-09-348-817A-4

Query Match 100.0%; Score 55; DB 4; Length 402;
Best Local Similarity 100.0%; Pred. No. 0.0085;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RMAPEIIMDR 11
Db 369 RMAPEIIMDR 379

RESULT 10
US-09-026-408-3
Sequence 3, Application US/09026408
Patent No. 6303338

GENERAL INFORMATION:

APPLICANT: NI et al.
TITLE OF INVENTION: PANCREAS-DERIVED PLASMINOGEN ACTIVATOR
TITLE OF INVENTION: INHIBITOR
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 100 NEW YORK AVENUE, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/026,408
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/934,011
FILING DATE: 15-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/024,056
FILING DATE: 16-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.0300002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 402 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-026-408-3

Query Match 100.0%; Score 55; DB 4; Length 402;
Best Local Similarity 100.0%; Pred. No. 0.0085;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPEITIMDR 11
DB 369 MAPEITIMDR 379

RESULT 11

US-08-781-020-6
Sequence 6, Application US/08781020
Patent No. 5792749
GENERAL INFORMATION:
APPLICANT: Wright, H. Tonle
TITLE OF INVENTION: METHOD AND COMPOSITION FOR LOWERING LOW
DENSITY LIPOPROTEIN CHOLESTEROL
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: WHITHAM, CURTIS, WHITHAM & MCGINN
STREET: Reston Intl. Center, 11800 Sunrise Valley Dr.,
CITY: Reston
STATE: VA
COUNTRY: USA
ZIP: 20191
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,020
FILING DATE: 09-JAN-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Whitham, Michael E.
REGISTRATION NUMBER: 32,635
REFERENCE/DOCKET NUMBER: VCU-16
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-391-2510
TELEFAX: 703-391-9035
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-781-020-6

Query Match 90.9%; Score 50; DB 1; Length 33;
Best Local Similarity 100.0%; Pred. No. 0.0049;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MAPEITIMDR 11
DB 1 MAPEITIMDR 10

RESULT 12

US-09-038-935-6
Sequence 6, Application US/09038935
Patent No. 6150332
GENERAL INFORMATION:
APPLICANT: Wright, H. Tonle
TITLE OF INVENTION: METHOD AND COMPOSITION FOR LOWERING LOW DENSITY
LIPOPROTEIN CHOLESTEROL
FILE REFERENCE: 2840223a
CURRENT APPLICATION NUMBER: US/09/038,935
CURRENT FILING DATE: 1998-03-12
EARLIER APPLICATION NUMBER: 08/781,020
EARLIER FILING DATE: 1997-01-19
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 6
LENGTH: 33
TYPE: PRT
ORGANISM: Homo sapiens
US-09-038-935-6

Query Match 90.9%; Score 50; DB 4; Length 33;
Best Local Similarity 100.0%; Pred. No. 0.0049;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MAPEITIMDR 11
DB 1 MAPEITIMDR 10

RESULT 13

US-08-041-774-3
Sequence 3, Application US/08041774
Patent No. 5550114
GENERAL INFORMATION:
APPLICANT: Strayer, David S.
TITLE OF INVENTION: EPIDERMAL GROWTH FACTOR INHIBITOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:

Tue Mar 19 11:00:02 2002

us-09-544-665-2.ra1

Page 6

```

ADDRESSER: COOPER & DUNHAM
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/041,774
FILING DATE: 19930402
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: ZIVIL, NO. 5550114man H.
REGISTRATION NUMBER: 25,385
REFERENCE/DOCKET NUMBER: 2560/42559
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 977-9809
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: N
ANTI-SENSE: N
FRAGMENT TYPE: Internal
US-08-041-774-3

Query Match          61.8%; Score 34; DB 1; Length 37;
Best Local Similarity 55.6%; Pred. No. 5.2;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 MAPEIIMD 10
DB 16 LAPEVYMD 24

RESULT 14
US-08-530-340-2
; Sequence 2, Application US/08530340A
; Patent No. 6255452
; GENERAL INFORMATION:
; APPLICANT: Strayer, David S
; TITLE OF INVENTION: Epidermal Growth Factor Inhibitor
; FILE REFERENCE: JEFF-0226
; CURRENT APPLICATION NUMBER: US/08/530,340A
; FILING DATE: 1995-12-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: An epidermal
; OTHER INFORMATION: growth factor inhibitor peptide
US-08-530-340-2

Query Match          61.8%; Score 34; DB 4; Length 37;
Best Local Similarity 55.6%; Pred. No. 5.2;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 MAPEIIMD 10
DB 16 LAPEVYMD 190
```

```

DB 16 LAPEVYMD 24

RESULT 15
US-08-041-774-5
; Sequence 5, Application US/08041774
; Patent No. 5550114
; GENERAL INFORMATION:
; APPLICANT: Strayer, David S.
; TITLE OF INVENTION: EPIDERMAL GROWTH FACTOR INHIBITOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: COOPER & DUNHAM
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
ZIP: 10112

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/041,774
FILING DATE: 19930402
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: ZIVIL, NO. 5550114man H.
REGISTRATION NUMBER: 25,385
REFERENCE/DOCKET NUMBER: 2560/42559
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 977-9809
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 216 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: N
ANTI-SENSE: N
FRAGMENT TYPE: Internal
US-08-041-774-5

Query Match          61.8%; Score 34; DB 1; Length 216;
Best Local Similarity 55.6%; Pred. No. 34;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 MAPEIIMD 10
DB 182 LAPEVYMD 190

Search completed: March 18, 2002, 09:49:40
Job time: 224 sec
```

Tue Mar 19 11:00:02 2002

us-09-544-665-2.rai

Page 7

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:50:28 ; Search time 41.98 Seconds
(without alignments)
19,960 Million cell updates/sec

Title: US-09-544-665-2

Perfect score: 55

Sequence: 1 RMAPEITMDR 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR-68:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	55	100.0	JC4265	plasminogen activa
2	55	100.0	ITRHUP1	plasminogen activa
3	55	100.0	S06745	plasminogen activa
4	42	76.4	A35032	plasminogen activa
5	41	74.5	S01240	heat shock protein
6	41	74.5	H85907	hypothetical prote
7	41	74.5	A34761	plasminogen activa
8	37	67.3	A86217	protein 123618.18
9	36	65.5	T32649	beta-1,2-xylosyltr
10	36	65.5	E88376	delta 3' to phbc -
11	35	63.6	D69274	medium-chain acyl-
12	35	63.6	S50543	GLN3 protein - yea
13	35	63.6	913	DNA-directed DNA p
14	35	63.6	S29529	genome polyprotein
15	34	61.8	48	A56771
16	34	61.8	170	A86073
17	34	61.8	170	S40802
18	34	61.8	176	B70429
19	34	61.8	216	TYHUC3
20	34	61.8	216	S24031
21	34	61.8	216	UC1455
22	34	61.8	216	GTP-binding protei
23	34	61.8	216	GTP-binding protei
24	34	61.8	217	gene Ran protein -
25	34	61.8	284	T44404
26	34	61.8	285	E84616
27	34	61.8	348	T29490
28	34	61.8	398	D72274
29	34	61.8	402	F82081
				cell division prot
				probable transamin

30	34	61.8	772	2	T08226	hypothetical prote
31	34	61.8	852	2	A75253	glutaminyl-tRNA sy
32	34	61.8	888	2	S61619	str2 protein - yea
33	34	61.8	941	2	T37626	DNA mismatch repa
34	33	60.0	128	2	A49166	3',5'-cyclic-AMP p
35	33	60.0	216	2	A70545	hypothetical prote
36	33	60.0	225	2	H75381	DNA-binding respon
37	33	60.0	268	2	T19236	hypothetical prote
38	33	60.0	268	2	S47087	PI7B protein - 11
39	33	60.0	292	2	F64486	hypothetical prote
40	33	60.0	395	2	JV0020	cystathionine beta
41	33	60.0	403	2	B83293	probable aminotran
42	33	60.0	403	2	T38078	hypothetical prote
43	33	60.0	416	2	T24564	hypothetical prote
44	33	60.0	489	2	T35375	probable membrane
45	33	60.0	492	2	F86384	probable protein f

ALIGNMENTS

RESULT 1
JC4265
plasminogen activator inhibitor type 1 precursor - American mink
C:Species: Mustela vison (American mink)
C:Date: 19-Oct-1995 #sequence_revision 08-Feb-1996 #text_change 20-Jun-2000
C:Accession: JC4265
R:Chuang, T.H.; Hamilton, R.T.; Nilsen-Hamilton, M.
Gene 162, 303-308, 1995
A:Title: Cloning of the mink plasminogen activator inhibitor type-1 messenger RNA. An
A:Reference number: JC4265; WID:96032362
A:Accession: JC4265
A:Molecule type: mRNA
A:Residues: 1-400 <CHU>
A:Cross-references: EMBL:X58541; NID:g1164923; PIDN:CAA41433.1; PID:g1164924
A:Experimental source: lung CCL64 epithelial cells
C:Comment: This protein controls the activities of the plasminogen activators and pla
C:Gene: pai-1
C:Superfamily: antithrombin III
C:Keywords: glycoprotein; plasminogen activator; serine proteinase inhibitor
F:1-21/Domain: signal sequence #status predicted <Sic> 1 #status predicted <MA>
F:22-400/Product: plasminogen activator inhibitor type 1 #status predicted <MA>
F:230-386/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:367/Inhibitory site: Arg (plasminogen activator) #status predicted

Query Match 100.0%; Score 55; DB 2; Length 400;
Best Local Similarity 100.0%; Pred. No. 0.0067;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RMAPEITMDR 11
DB 367 RMAPEITMDR 377

RESULT 2
ITRHUP1
plasminogen activator inhibitor 1 precursor [validated] - human
N:Alternate names: plasminogen activator inhibitor, endothelial
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 08-Dec-2000
C:Accession: A28107; S02551; A26996; I59126; J50397; A25693; A26146; A29100; A25895;
R:Bosma, P.J.; van den Berg, E.A.; Koolstra, T.; Stenlehik, D.R.; Slighem, J.L.
J. Biol. Chem. 263, 9129-9141, 1988
A:Title: Human plasminogen activator inhibitor-1 gene. Promoter and structural gene n
A:Reference number: A28107; WID:88243790
A:Accession: A28107
A:Molecule type: DNA
A:Residues: 1-402 <BOS>
A:Cross-references: GB:J03764; NID:g189564; PIDN:AAA60007.1; PID:g386996
R:Strandberg, L.; Lawrence, D.; NY, T.
Eur. J. Biochem. 176, 609-616, 1988

A>Title: The organization of the human-plasminogen-activator-inhibitor-1 gene. Implicat
A:Reference number: S02551; MUID:89005111
A:Accession: S02551
A:Molecule type: DNA
A:Residues: 1-14, 'T', 16-402 <STR>
A:Cross-references: EMBL:X1338; NID:g35244; PIDN:CA28442.1; PID:g35245
A:Note: the complete translation is not annotated in Genbank entries HSPAI1, HSPAI12, H
R:Loskutoff, D.J.; Linders, M.; Keltjer, J.; Veerman, H.; van Heerikhuizen, H.; Pannekoek
Blochmestry 26, 3763-3768, 1987
A>Title: Structure of the human plasminogen activator inhibitor 1 gene: nonrandom distr
A:Reference number: A26996; MUID:88000586
A:Accession: A26996
A:Molecule type: DNA
A:Residues: 1-402 <LOS>
A:Cross-references: GB:M2321; GB:M17121; NID:g189576; PIDN:AAA60009.1; PID:g189578
A:Note: the sequence in Genbank entry H0MPA182, release 109.0, (PID:g189578) has the cod
R:van Zonneveld, A.J.; Curriden, S.A.; Loskutoff, D.J.
Proc Natl. Acad. Sci. U.S.A. 85, 5525-5529, 1988
A>Title: Type I plasminogen activator inhibitor gene: functional analysis and glucocorti
A:Reference number: I59126; MUID:88289734
A:Accession: I59126
A:Molecule type: DNA
A>Status: translated from GB/EMBL/DBJ
A:Residues: 1-9 <ZON>
A:Cross-references: GB:J03836; NID:g189579; PIDN:AAA60010.1; PID:g189580
R:Follo, M.; Ginsburg, D.
Gene 84, 447-453, 1989
A>Title: Structure and expression of the human gene encoding plasminogen activator inh
A:Reference number: J50397; MUID:50128289
A:Accession: J50397
A:Molecule type: DNA
A:Residues: 85-86; 88-93; 166-171; 231-236; 298-302; 331-336; 360-365; 388-393 <FOL>
A:Cross-references: GB:M3136; NID:g189543
A:Note: sequences of the intron/exon boundaries are shown
R:Pannekoek, H.; Veerman, H.; Lambers, H.; Diergaarde, P.; Verweij, C.L.; van Zonneveld,
EMBO J. 5, 2539-2544, 1986
A>Title: Endothelial plasminogen activator inhibitor (PAI): a new member of the serpin g
A:Reference number: A91052; MUID:87053819
A:Accession: A25693
A:Molecule type: mRNA
A:Residues: 1-402 <PAN>
A:Cross-references: GB:X04429; NID:g35271; PIDN:CA28025.1; PID:g35272
R:Ginsburg, D.; Zehed, R.; Yang, A.Y.; Rafferty, U.M.; Andreasen, P.A.; Nielsen, L.; Dan
J. Clin. Invest. 78, 1673-1680, 1986
A>Title: cDNA cloning of human plasminogen activator-inhibitor from endothelial cells.
A:Reference number: A92766; MUID:87058123
A:Accession: A26146
A:Molecule type: mRNA
A:Residues: 1-402 <GIN>
A:Cross-references: GB:M16006; NID:g189541; PIDN:AAA60003.1; PID:g189542
R:Shun, T.C.; Kretzmer, K.K.
FEBS Lett. 210, 11-16, 1987
A>Title: cDNA cloning and expression in E. coli of a plasminogen activator inhibitor (PA
A:Reference number: A29100; MUID:87105925
A:Accession: A29100
A:Molecule type: mRNA
A:Residues: 17-402 <MUN>
A:Cross-references: GB:X04744; NID:g35275; PIDN:CA28444.1; PID:g352747
A:Note: part of this sequence, including the amino end of the mature protein, was confir
R:Wu, T.; Sawley, M.; Lawrence, D.; Millan, J.L.; Loskutoff, D.J.
Proc. Natl. Acad. Sci. U.S.A. 83, 6776-6780, 1986
A>Title: Cloning and sequence of a cDNA coding for the human beta-migrating endothelial
A:Reference number: A25895; MUID:86313660
A:Accession: A25895
A:Molecule type: mRNA
A:Residues: 20-402 <NVT>
A:Cross-references: GB:M14083; NID:g189566; PIDN:AAA60008.1; PID:g386997
R:Andreasen, P.A.; Riccio, A.; Wellinder, K.G.; Douglas, R.; Sartorio, R.; Nielsen, L.S.;
FEBS Lett. 209, 213-218, 1986
A>Title: Plasminogen activator inhibitor type-1: reactive center and amino-terminal hept
A:Reference number: A91371; MUID:87080762
A:Accession: A25651
A:Molecule type: mRNA

A:Residues: 1-14, 'T', 16-47 <AND1>
A:Cross-references: GB:X04729; NID:g35263; PIDN:CA28438.1; PID:g35264
A:Accession: B25651
A:Molecule type: mRNA
A:Residues: 364-402 <AND2>
A:Cross-references: GB:X04731; NID:g35260; PIDN:CA28442.1; PID:g35261
R:Lang, W.E.; Abernethy, R.; Jong, A.; Rideout, W.; Bergman, B.L.; Baker, J.
Thromb. Haemost. 61, 517-521, 1989
A>Title: Isolation of multiple types of plasminogen activator inhibitors from vascula
A:Reference number: A60436; MUID:90020174
A:Accession: A60436
A:Molecule type: protein
A:Residues: 225-235 <LAU>
R:Kjoller, L.; Mathiesen, P.M.; Sottorp-Jensen, L.; Justesen, J.; Rodenburg, K.W.; An
Eur. J. Biochem. 241, 38-46, 1996
A>Title: Conformational changes of the reactive-centre loop and beta-strand 5A accomp
A:Reference number: S74133; MUID:97054589
A:Accession: S74133
A:Molecule type: protein
A:Residues: 22-30; 370-376 <KIO>
R:Stromqvist, M.; Karlsson, K.E.; Bjorquist, P.; Andersson, J.O.; Bystrom, M.; Han
Biochim. Biophys. Acta 1293, 103-109, 1996
A>Title: Characterization of the complex of plasminogen activator inhibitor type 1 wi
A:Reference number: S70346; MUID:96283799
A:Accession: S70346
A>Status: preliminary
A:Molecule type: protein
A:Residues: 370-375 <SNO>
C:Comment: This inhibitor acts as "bait" for tissue plasminogen activator (see PIR:UK
fibrinolysis. High concentrations of this protein have been associated with human th
C:Comment: Three types of PAI have been identified (see also PIR:A32853 and PIR:A3933
oma cells. Vascular endothelial cells may be the primary site of synthesis of plasma
C:Comment: Glycosylation is not required for inhibitory activity.
C:Genetics:
A:Gene: GDB:PAI1; PLANH1
A:Cross-references: GDB:120297; OMIM:173360
A:Map position: 7q21.3-7q22
A:Insertions: 91/1; 169/1; 234/1; 300/2; 334/1; 363/1; 391/1
C:Superfamily: antithrombin III
C:Keywords: glycoprotein; serine proteinase inhibitor
F:1-23/Domain: signal sequence status predicted <SIG>
F:24-402/Product: plasminogen activator inhibitor-1 status experimental <MNT>
F:222,286,352/Binding site: carbohydrate (Asn) (covalent) status predicted
F:369/Inhibitory site: Arg (plasminogen activator) status predicted

Query Match 100.0%; Score 55; DB 1; Length 402;
Best Local Similarity 100.0%; Pred. No. 0.0067;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RMAPEIIMDR 11
DB 369 RMAPEIIMDR 379

RESULT 3
506745
plasminogen activator inhibitor-1 precursor - bovine
N:Alternate names: endothelial-cell plasminogen activator inhibitor; PAI-1
C:Species: Bos primigenius taurus (cattle)
C>Date: 28-Feb-1990 #sequence revision 22-Apr-1995 #text_change 18-Jun-1999
C:Accession: S06745; #35855; S01324; S10906
R:Mimuro, J.; Sawley, M.; Hattori, M.; Loskutoff, D.J.
Nucleic Acids Res. 17, 8872, 1989
A>Title: cDNA for bovine type I plasminogen activator inhibitor (PAI-1).
A:Reference number: S06745; MUID:90067867
A:Accession: S06745
A:Molecule type: mRNA
A:Residues: 1-802 <MTM>
A:Cross-references: EMBL:X16383; NID:9600; PIDN:CA24419.1; PID:g601
R:Pepper, M.S.; Bellin, D.; Montesano, R.; Orci, L.; Vassalli, J.D.
J. Cell Biol. 111, 743-755, 1990
A>Title: Transforming growth factor-beta 1 modulates basic fibroblast growth factor-1

A:Reference number: A35855; MUID:90338128
 A:Accession: A35855
 A:Molecule type: mRNA
 A:Residues: 153-235 <PEP>
 A:Cross-references: EMBL:X52906; NID:g598; PIDN:CAA37094.1; PID:g930005
 R:Kabatgiri, K.; Okada, K.; Hattori, H.; Iano, M.
 Eur. J. Biochem. 176, 81-87, 1988
 A:Title: Bovine endothelial cell plasminogen activator inhibitor. Purification and heat
 A:Reference number: S01324; MUID:88329072
 A:Accession: S01324
 A:Molecule type: protein
 A:Residues: 24-49; 'L', '51-63 <KAT>
 C:Comment: Three types of PAI have been identified. PAI-1 is an acid-stable glycoprotein
 may be the primary site of synthesis of plasma PAI.
 C:Comment: This inhibitor acts as "bait" for tissue plasminogen activator, urokinase, an
 yals.
 C:Comment: Glycosylation is not a prerequisite for inhibitory activity.
 C:Superfamily: antithrombin III
 C:Keywords: glycoprotein; serine protease inhibitor
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-402/Product: plasminogen activator inhibitor 1 #status predicted <MAT>
 F:232,288,352/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:369/Inhibitory site: Arg (plasminogen activator) #status predicted

Query Match 100.0%; Score 55; DB 1; Length 402;
 Best Local Similarity 100.0%; Pred. No. 0.0067;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RMAPEITIMDR 11
 |||||
 DB 369 RMAPEITIMDR 379

RESULT 4
 A35032
 plasminogen activator inhibitor 1 precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 10-Sep-1999 #sequence-revision 10-Sep-1999 #text-change 10-Sep-1999
 C:Accession: A35032; J0490; A60581; A39120
 R:Brizdinski, C.J.; Riordan-Johnson, M.; Nordby, E.C.; Suter, S.M.; Gelehrter, T.D.
 J. Biol. Chem. 265, 2078-2085, 1990
 A:Title: Isolation and characterization of the rat plasminogen activator inhibitor-1 gen
 A:Reference number: A35032; MUID:90130456
 A:Accession: A35032
 A:Molecule type: DNA
 A:Residues: 1-402 <BRD>
 A:Cross-references: GB:J05206; NID:g205965; PIDN:AAA1796.1; PID:g205966
 R:Zeheb, R.; Gelehrter, T.D.
 Gene 73, 459-468, 1988
 A:Title: Cloning and sequencing of cDNA for the rat plasminogen activator inhibitor-1.
 A:Reference number: J0490; MUID:89211983
 A:Accession: J0490
 A:Molecule type: mRNA
 A:Residues: 1-402 <ZEB>
 A:Cross-references: GB:M24067; NID:g577500; PIDN:AAA56856.1; PID:g577501
 R:Newman, M.J.; Lane, E.A.; Iannotti, A.M.; Nugent, M.A.; Pehlinsky, R.B.; Keski-Oja, J.
 Endocrinology 126, 2936-2946, 1990
 A:Title: Characterization and purification of a secreted plasminogen activator inhibitor
 on in transformed NRK cells.
 A:Reference number: A60581; MUID:90276328
 A:Accession: A60581
 A:Molecule type: protein
 A:Residues: 24-48 <ZEB>
 R:Olson Jr., J.A.; Shaverick, K.T.; Ogilvie, S.; Buhl, W.C.; Raizada, M.K.
 Proc. Natl. Acad. Sci. U.S.A. 88, 1928-1932, 1991
 A:Title: Angiotensin II induces secretion of plasminogen activator inhibitor 1 and a tis
 A:Reference number: A39120; MUID:91156719
 A:Accession: A39120
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 24-43; 'G' <OLS>
 C:Genetics:

A:introns: 91/1; 169/1; 234/1; 300/2; 334/1; 363/1; 391/1
 C:Superfamily: antithrombin III
 C:Keywords: glycoprotein
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-402/Product: plasminogen activator inhibitor-1 #status experimental <MAT>
 F:68,232,288,352/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:369/Inhibitory site: Arg (plasminogen activator) #status predicted

Query Match 76.4%; Score 42; DB 1; Length 402;
 Best Local Similarity 63.6%; Pred. No. 2.2;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 RMAPEITIMDR 11
 |||||
 DB 369 RMAPEITIMDR 379

RESULT 5
 S01240
 heat shock protein grpe (heat shock protein b25.3) (hsp24) - Escherichia coli
 N:Alternate names: grpe protein
 C:Species: Escherichia coli
 C:Date: 28-Aug-1989 #sequence-revision 28-Aug-1989 #text-change 24-Sep-1999
 C:Accession: S01240; A65040
 R:Lipinska, B.; King, J.; Ang, D.; Georgopoulos, C.
 Nucleic Acids Res. 16, 7545-7562, 1988
 A:Title: Sequence analysis and transcriptional regulation of the Escherichia coli gr
 A:Reference number: S01240; MUID:88319942
 A:Accession: S01240
 A:Molecule type: DNA
 A:Residues: 1-197 <LIP>
 A:Cross-references: EMBL:X07863; NID:g41618; PIDN:CAA30711.1; PID:g41619
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.
 .A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MUID:97426617
 A:Accession: A65040
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-197 <BLAT>
 A:Cross-references: GB:AE000347; GB:U00096; NID:g2367142; PIDN:NACT5663.1; PID:g1788
 A:Experimental source: strain K-12, substrain MG1655
 C:Genetics:
 A:Gene: grpe
 C:Superfamily: heat shock protein grpe
 F:2-197/Product: heat shock protein B25.3 #status predicted <MAT>

Query Match 74.5%; Score 41; DB 2; Length 197;
 Best Local Similarity 88.9%; Pred. No. 1.6;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 APEITIMDR 11
 |||||
 DB 13 APEITIMDR 21

RESULT 6
 H85907
 hypothetical protein grpe [Imported] - Escherichia coli (strain O157:H7)
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence-revision 16-Feb-2001 #text-change 31-Mar-2001
 C:Accession: H85907
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Ma
 iller, L.; Grobeck, E.J.; Davis, N.W.; Llin, A.; Diallanta, E.; Potamoukis, K.; Apod
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: H85907
 A:Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-197 <STO>
 A:Cross-references: GB:AE005174; NID:g12517038; PIDN:AAC57724.1; GSPDB:GN00145; UWGCP:239
 A:Experimental source: strain 0157:H7, substrain EDL933
 C:Genetics:
 A:Gene: gripe
 C:Superfamily: heat shock protein gripe

Query Match 74.5%; Score 41; DB 2; Length 197;
 Best Local Similarity 88.9%; Pred. No. 1.6;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 APEEIIIMDR 11
 |||||
 DB 13 APEEIIIMDQ 21

RESULT 7
 A34761
 plasmidogen activator inhibitor 1 homolog mrl - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: A34761; S48208
 R:Prendergast, G.C.; Diamond, L.E.; Dahl, D.; Cole, M.D.
 Mol. Cell. Biol. 10, 1265-1269, 1990
 A:Title: The c-myc-regulated gene mrl encodes plasmidogen activator inhibitor 1.
 A:Reference number: A34761; MUID:90156593
 A:Accession: A34761
 A:Molecule type: mRNA
 A:Residues: 1-402 <PRE>
 A:Cross-references: GB:M33960; NID:g200219; PIDN:AAA39887.1; PID:g200220
 R:Liljen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.
 Eur. J. Biochem. 224, 863-871, 1994
 A:Title: Characterization of the murine plasma fibrinolytic system.
 A:Reference number: S48202; MUID:95010076
 A:Accession: S48208
 A:Molecule type: protein
 A:Residues: 'M', 24-29 <LIJ>
 C:Superfamily: antithrombin III

Query Match 74.5%; Score 41; DB 1; Length 402;
 Best Local Similarity 63.6%; Pred. No. 3.5;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 RMAPEEIIIMDR 11
 |||||
 DB 369 RMAPEEIIIMDR 379

RESULT 8
 AB6217
 Protein T23G18.18 [Imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: AB6217
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Child, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 Jensen, N.F.; Hughes, B.; Hultzer, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaytin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Malt, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: AB6141; MUID:21016719
 A:Accession: AB6217
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-382 <STO>
 A:Cross-references: GB:AE005172; NID:g6579201; PIDN:AAF18244.1; GSPDB:GN00141
 C:Genetics:

A:Gene: T23G18.18
 A:Map position: 1

Query Match 67.3%; Score 37; DB 2; Length 382;
 Best Local Similarity 70.0%; Pred. No. 20;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 RMAPEEIIIMDR 11
 |||||
 DB 345 IAPPEEIIIMDR 354

RESULT 9
 T52649
 beta-1,2-xylosyltransferase [validated] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 02-Mar-2001
 C:Accession: T52649
 R:Strasser, R.; Mucha, J.; Mach, L.; Altmann, F.; Wilson, I.B.H.; Gloessel, J.; Stein
 FEBS Lett. 472, 105-108, 2000
 A:Title: Molecular cloning and functional expression of beta 1,2-xylosyltransferase
 A:Reference number: T26156
 A:Accession: T52649
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-534 <STR>
 A:Cross-references: EMBL:AJ272121; PIDN:CAB90610.1
 C:Genetics:
 A:Gene: xy1T
 C:Function:
 A:Description: catalyzes the transfer of xylose from UDP-xylose to the core beta-11n)
 C:Superfamily: Arabidopsis thaliana beta-1,2-xylosyltransferase
 C:Keywords: glycosyltransferase

Query Match 65.5%; Score 36; DB 2; Length 534;
 Best Local Similarity 54.5%; Pred. No. 45;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 RMAPEEIIIMDR 11
 |||||
 DB 127 RMAPEEIIIMDR 137

RESULT 10
 E48376
 orfs 3' to phbc - Thioecystis violacea
 C:Species: Thioecystis violacea
 C:Date: 19-Nov-1993 #sequence_revision 25-Apr-1997 #text_change 21-Jul-2000
 C:Accession: E48376
 R:Liebergesell, M.; Steinbuechel, A.
 Appl. Microbiol. Biotechnol. 38, 493-501, 1993
 A:Title: Cloning and molecular analysis of the poly(3-hydroxybutyric acid) biosynthes
 A:Reference number: A48376; MUID:93159750
 A:Accession: E48376
 A:Status: preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-692 <LIJ>
 A:Cross-references: GB:554369; NID:g298249; PIDN:AAC60431.1; PID:g298254
 A:Note: sequence extracted from NCBI backbone (NCBIN:124660, NCBI:P:124665)

Query Match 65.5%; Score 36; DB 2; Length 692;
 Best Local Similarity 60.0%; Pred. No. 60;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 RMAPEEIIIMDR 10
 |||||
 DB 49 RMAPEEIIIMDR 58

RESULT 11

D69274
 medium-chain acyl-CoA ligase (alk-1) homolog - Archaeoglobus fulgidus
 C:Species: Archaeoglobus fulgidus
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 18-Aug-2000
 C:Accession: D69274
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
 .; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
 .; Glodek, A.; Zhou, L.; Overbeek, K.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 .; Natune, J.O.; White, C.R.; Venter, J.C.
 A:Authors: Overbeek, K.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 .; Smith, H.O.; Moese, C.R.; Venter, J.C.
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae
 A:Reference number: A69250; MUID:98049343
 A:Accession: D69274
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1540 <KIR>
 A:Cross-references: GB:AE001093; GB:AE000782; NID:92689416; PIDN:AB91037.1; PID:9265044
 C:Superfamily: probable acyl-CoA ligase medium chain; acetate--CoA ligase homology
 F:50-530/Domain: acetate--CoA ligase homology <ACLD>

Query Match 63.6%; Score 35; DB 2; Length 540;
 Best Local Similarity 75.0%; Pred. No. 71;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 RMAPEIIM 8
 DB 87 RMAPEIIV 94

RESULT 12
 S50543
 GLN3 protein - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein YER040w
 C:Species: Saccharomyces cerevisiae
 C>Date: 28-Jan-1995 #sequence_revision 12-May-1995 #text_change 29-Oct-1999
 C:Accession: S50543; S22280
 R:Dieterich, P.S.
 submitted to the EMBL Data Library, December 1994
 A:Description: The sequence of S. cerevisiae cosmid 9379, 9581, and lambda clone 4678.
 A:Reference number: S50536
 A:Accession: S50543
 A:Molecule type: DNA
 A:Residues: 1-730 <DIE>
 A:Cross-references: EMBL:U18796; NID:9603265; PIDN:AB64575.1; PID:9603273; MIPS:YER040w
 R:Minehart, P.L.; Magasanik, B.
 Mol. Cell. Biol. 11, 6216-6228, 1991
 A:Title: Sequence and expression of GLN3, a positive nitrogen regulatory gene of Sacchar
 A:Reference number: S22280; MUID:92049353
 A:Accession: S22280
 A:Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-473, G', 475-730 <MIN>
 C:Genetics:
 A:Gene: SGD:GLN3
 A:Cross-references: SGD:S0000842; MIPS:YER040w
 A:Map position: 5R
 C:Superfamily: GATA-type zinc finger homology
 C:Keywords: zinc finger
 F:303-356/Domain: GATA-type zinc finger homology <GZF>
 F:306-330/Region: zinc finger GATA motif

Query Match 63.6%; Score 35; DB 2; Length 730;
 Best Local Similarity 70.0%; Pred. No. 99;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 RMAPEIIMD 10
 DB 403 RVPEIIR 412

RESULT 13
 B42543
 DNA-directed DNA polymerase (EC 2.7.7.7) - Chlorella virus CV-NY-2A
 C:Species: Chlorella virus CV-NY-2A
 C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 11-Jun-1999
 C:Accession: B42543
 R:Grabherr, R.; Strasser, P.; Van Etten, J.L.
 Virology 188, 721-731, 1992
 A:Title: The DNA polymerase gene from chlorella viruses PBCV-1 and NY-2A contains an
 A:Reference number: A42543; MUID:92263776
 A:Accession: B42543
 A:Molecule type: DNA
 A:Residues: 1-913 <GRA>
 A:Cross-references: GB:M86837; NID:9323321; PIDN:AAA8827.1; PID:9323322
 C:Genetics:
 A:Insertions: 682/3
 C:Superfamily: hepesvirus DNA-directed DNA polymerase
 C:Keywords: DNA binding; DNA biosynthesis; DNA replication; exonuclease; nucleotidyl
 F:182-199/Region: exonuclease pattern A
 F:268-282/Region: exonuclease pattern B
 F:388-401/Region: exonuclease pattern C

Query Match 63.6%; Score 35; DB 1; Length 913;
 Best Local Similarity 50.0%; Pred. No. 1,3e+02;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 2 NAPEIIMDR 11
 DB 499 MSPETIVMDK 508

RESULT 14
 S29529
 genome polyprotein - Toscana virus
 N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)
 C:Species: Toscana virus
 C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 24-Nov-1999
 C:Accession: S29529
 R:Accardi, L.; Gro, M.C.; di Bonito, P.; Giorgi, C.
 submitted to the EMBL Data Library, September 1992
 A:Description: Toscana virus genomic L segment: molecular cloning, coding strategy a
 A:Reference number: S29529
 A:Accession: S29529
 A:Molecule type: genomic RNA
 A:Residues: 1-2095 <ACC>
 A:Cross-references: EMBL:X68414; NID:962199; PIDN:CAA48478.1; PID:962200
 C:Superfamily: unsegmented virus RNA-directed RNA polymerase
 C:Keywords: nucleotidyltransferase

Query Match 63.6%; Score 35; DB 2; Length 2095;
 Best Local Similarity 45.5%; Pred. No. 3.2e+02;
 Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 1 RMAPEIIMDR 11
 DB 1519 RVAPKIVSDK 1529

RESULT 15
 A56771
 34K epidermal growth factor inhibitor, ras-like - rabbit (fragments)
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C>Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 28-May-1999
 C:Accession: A56771
 R:Strayer, D.S.; Mathew, J.
 Am. J. Pathol. 142, 1141-1153, 1993
 A:Title: A 34-kd protein with strong homology to ras-like proteins inhibits epiderma
 A:Reference number: A56771; MUID:93235943
 A:Accession: A56771
 A:Status: preliminary
 A:Molecule type: protein

A:Residues: 1-48 <STR>
C:Superfamily: ras transforming protein; translation elongation factor Tu homology
C:Keywords: GTP binding

Query Match 61.8% Score 34; DB 2; Length 48;
Best Local Similarity 55.6%; Pred No 7.4;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 2 MAPPEIIMD 10
: 1 1 1 1 1
DB 27 LAPEEYVMD 35

Search completed: March 18, 2002, 09:50:30
Job time: 239 sec

OY 1 RMAPEEIMDR 11
 DB 367 RMAPEEIMDR 377

RESULT 2
 PAI1_BOVIN STANDARD: PRT: 402 AA.

ID PAI1_BOVIN STANDARD: PRT: 402 AA.

AC P13909;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE PLASMINOGEN ACTIVATOR INHIBITOR-1 PRECURSOR (PAI-1) (ENDOTHELIAL
 DE PLASMINOGEN ACTIVATOR INHIBITOR) (PAI).
 GN SERPINE1 OR PAI1.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_Taxid:9913;
 OX NCBI_Taxid:9913;
 RN 1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE-90067867; PubMed-2587231; Loskutoff D.J.;
 RA Mamo J., Sawdey M., Hattori M., Loskutoff D.J.;
 RT "CDNA for bovine type 1 plasminogen activator inhibitor (PAI-1).";
 RL Nucleic Acids Res. 17:8872-8872(1989).
 RN 2)
 RP SEQUENCE OF 24-63.
 RX MEDLINE-88329072; PubMed-3262060;
 RA Katsagiri K., Okada K., Hattori H., Yano M.;
 RT "Bovine endothelial cell plasminogen activator inhibitor.
 RL Purification and heat activation.";
 RL Eur. J. Biochem. 176:81-87(1988).
 RN 3)
 RP SEQUENCE OF 153-235 FROM N.A.
 RC TISSUE-ADRENAL CORTEX;
 RX MEDLINE-9038128; PubMed-1696269;
 RA Pepper M.S., Belin D., Montesano R., Orci L., Vassalli J.-D.;
 RT "Transforming growth factor-beta 1 modulates basic fibroblast growth
 factor-induced proteolytic and angiogenic properties of endothelial
 cells in vitro.";
 RL J. Cell Biol. 111:743-755(1990).
 RN 4)
 RP FUNCTION: THIS INHIBITOR ACTS AS "PAI" FOR TISSUE PLASMINOGEN
 ACTIVATOR, UROKINASE, AND PROTEIN C. ITS RAPID INTERACTION WITH
 TPA MAY FUNCTION AS A MAJOR CONTROL POINT IN THE REGULATION OF
 FIBRINOLYSIS.
 CC MISCELLANEOUS: PAI1 IS INACTIVATED BY PROTEOLYTIC ATTACK OF THE
 CC UROKINASE-TYPE (U-PA) AND THE TISSUE-TYPE (TPA), CLEAVING THE
 CC 369(R)-370(M) BOND.
 CC MISCELLANEOUS: VASCULAR ENDOTHELIAL CELLS MAY BE THE PRIMARY SITE
 CC OF SYNTHESIS OF PLASMA PAI.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
 CC THIS SWISS-PROT ENTRY IS COPYRIGHTED. IT IS PRODUCED THROUGH A COLLABORATION
 CC BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION
 CC AT THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS
 CC USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
 CC MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
 CC ENTITIES REQUIRES A LICENSE AGREEMENT (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X16383; CA34419.1;
 DR EMBL: X52906; CA37094.1;
 DR PIR: S01324; S01324;
 DR PIR: S06745; S06745;
 DR HSSP: P01008; IATH.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin.1.
 DR SMART: SM00093; SERPIN.1.
 DR PROSITE: PS00284; SERPIN.1.
 KW Serpin; Serine protease inhibitor; Plasma; Plasminogen activation;
 KW Glycoprotein; Signal.

FT SIGNAL 1 23
 FT CHAIN 24 402
 FT ACT_SITE 369 370
 FT CARBOHYD 232 232
 FT CARBOHYD 288 288
 FT CARBOHYD 352 352
 FT CONFLICT 50 50
 FT SEQUENCE 402 AA; 45371 MW; 9053617333C7D130 CRC64;

Query Match 100.0%; Score 55; DB 1; Length 402;
 Best Local Similarity 100.0%; Preg. No. 0.0017;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RMAPEEIMDR 11
 DB 369 RMAPEEIMDR 379

RESULT 3
 PAI1_HUMAN STANDARD: PRT: 402 AA.

ID PAI1_HUMAN STANDARD: PRT: 402 AA.

AC P05121;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE PLASMINOGEN ACTIVATOR INHIBITOR-1 PRECURSOR (PAI-1) (ENDOTHELIAL
 DE PLASMINOGEN ACTIVATOR INHIBITOR) (PAI).
 GN SERPINE1 OR PAI1 OR PLANH1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_Taxid:9606;
 OX NCBI_Taxid:9606;
 RN 1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE-87053819; PubMed-2430793;
 RA Pannekoek H., Veerman H., Lambers H., Diergaarde P., Verweij C.L.,
 RA van Zonneveld A.-J., van Mourik J.A.;
 RT "Endothelial plasminogen activator inhibitor (PAI): a new member of
 RL the Serpin gene family.";
 RL EMBO J. 5:2539-2544(1986).
 RN 2)
 RP SEQUENCE FROM N.A.
 RX MEDLINE-8800586; PubMed-2820474;
 RA Loskutoff D.J., Linders M., Kelder J., Veerman H.,
 RA van Heerikhuizen H., Pannekoek H.;
 RT "Structure of the human plasminogen activator inhibitor 1 gene:
 RL nonrandom distribution of introns.";
 RL Biochemistry 26:3763-3768(1987).
 RN 3)
 RP SEQUENCE FROM N.A.
 RX MEDLINE-87058123; PubMed-3097076;
 RA Ginsburg D., Zehnb R., Yang A.Y., Rafferty U.M., Andreassen P.A.,
 RA Nielsen L., Dano K., Lebo R.V., Gelehrter T.D.;
 RT "cDNA cloning of human plasminogen activator-inhibitor from
 RL endothelial cells.";
 RL J. Clin. Invest. 78:1673-1680(1986).
 RN 4)
 RP SEQUENCE FROM N.A.
 RX MEDLINE-90128289; PubMed-2612914;
 RA Folio M., Ginsburg D.;
 RT "Structure and expression of the human gene encoding plasminogen
 RL activator inhibitor, PAI-1.";
 RL Gene 84:447-453(1989).
 RN 5)
 RP SEQUENCE FROM N.A.
 RX MEDLINE-89005111; PubMed-3262512;
 RA Strandberg L., Lawrence D., Ny T.;
 RT "The organization of the human-plasminogen-activator-inhibitor-1
 RL gene. Implications on the evolution of the serine-protease inhibitor
 RL family.";
 RL Eur. J. Biochem. 176:609-616(1988).
 RN 6)

RP SEQUENCE FROM N.A.
 RX MEDLINE-88243790; PubMed-3132455;
 RA Bosma P.J.; Van den Berg E.A., Kooistra T., Stenliak D.R.,
 RA Slighom J.L.;
 RT "Human plasminogen activator inhibitor-1 gene. Promoter and
 RT structural gene nucleotide sequences.";
 RL J. Biol. Chem. 263:9129-9141(1988).
 RN
 RP SEQUENCE FROM N.A.
 RA Pannekoek H.;
 RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
 RN
 RP SEQUENCE FROM N.A.
 RA Cordes M., Doela D.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN
 RP SEQUENCE OF 20-402 FROM N.A.
 RX MEDLINE-86313660; PubMed-3092219;
 RA Ny T., Sawley M., Lawrence D., Millan J.L., Loskutov D.J.;
 RT "Cloning and sequence of a cDNA coding for the human beta-migrating
 RT endothelial-cell-type plasminogen activator inhibitor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:6776-6780(1986).
 RN
 RP SEQUENCE OF 1-47 AND 364-402 FROM N.A.
 RX MEDLINE-87080762; PubMed-3025016;
 RA Andreassen P.A., Riccio A., Wellinder K.G., Douglas R., Sartorio R.,
 RA Nielsen L.S., Oppenheimer C., Blasi F., Danco K.;
 RT "Plasminogen activator inhibitor type-1: reactive center and amino-
 RT terminal heterogeneity determined by protein and cDNA sequencing.";
 RL FEBS Lett. 209:213-218(1986).
 RN
 RP SEQUENCE OF 17-402 FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE-Placenta;
 RX MEDLINE-87105925; PubMed-3026837;
 RA Man T.C., Kretzmer K.K.;
 RT "cDNA cloning and expression in E. coli of a plasminogen activator
 RT inhibitor (PAI) related to a PAI produced by Hep G2 hepatoma cell.";
 RL FEBS Lett. 210:11-16(1987).
 RN
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
 RX MEDLINE-92114970; PubMed-1731226;
 RA Mottenson J., Strand A., Symesky J., Sweet R.M., Danley D.E.,
 RA Geoghegan K.F., Gerard R.D., Goldsmith E.J.;
 RT "Structural basis of latency in plasminogen activator inhibitor-1.";
 RL Nature 355:270-273(1992).
 RN
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
 RX MEDLINE-96003732; PubMed-7552714;
 RA Aertgeerts K., de Bondt H.L., de Ranter C.J., Declercq P.J.;
 RT "Mechanisms contributing to the conformational and functional
 RT flexibility of plasminogen activator inhibitor-1.";
 RL Nat. Struct. Biol. 2:891-897(1995).
 RN
 RP X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS).
 RX MEDLINE-98298439; PubMed-9634700;
 RA Xue Y., Bjorkquist P., Ingmarct T., Linschoten M., Musil D., Sjolén L.,
 RA Delnum J.;
 RT "Interfering with the inhibitory mechanism of serpins: crystal
 RT structure of a complex formed between cleaved plasminogen activator
 RT inhibitor type 1 and a reactive-centre loop peptide.";
 RL Structure 6:627-636(1998).
 RN
 RP X-RAY CRYSTALLOGRAPHY (2.99 ANGSTROMS).
 RX MEDLINE-99148119; PubMed-10368279;
 RA Sharp A.M., Stein P.E., Pannu N.S., Carrell R.W., Berkenpas M.B.,
 RA Ginsburg D., Lawrence D.A., Read R.J.;
 RT "The active conformation of plasminogen activator inhibitor 1, a
 RT target for drugs to control fibrinolysis and cell adhesion.";
 RL Structure 7:111-118(1999).
 RN
 RP VARIANT THR-15.
 RX MEDLINE-97337920; PubMed-9194591;
 RA Turkmen B., Schmitt M., Schmalheidt B., Trommler P., Hell W.,

RA Creutzburg S., Graeff H., Magdalen V.;
 RT "Mutational analysis of the genes encoding urokinase-type plasminogen
 RT activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer.";
 RL Electrophoresis 18:686-689(1997).
 CC
 CC -1- FUNCTION: THIS INHIBITOR ACTS AS "BAIT" FOR TISSUE PLASMINOGEN
 CC ACTIVATOR, UROKINASE, AND PROTEIN C. ITS RAPID INTERACTION WITH
 CC TPA MAY FUNCTION AS A MAJOR CONTROL POINT IN THE REGULATION OF
 CC FIBRINOLYSIS.
 CC
 CC -1- DISEASE: HIGH CONCENTRATIONS OF THIS PROTEIN HAVE BEEN ASSOCIATED
 CC WITH HUMAN THROMBOEMBOLIC DISEASE.
 CC
 CC -1- MISCELLANEOUS: PAI1 IS INACTIVATED BY PROTEOLYTIC ATTACK OF THE
 CC UROKINASE-TYPE (U-PA) AND THE TISSUE-TYPE (TPA), CLEAVING THE
 CC 369(R)-370(M) BOND.
 CC
 CC -1- MISCELLANEOUS: TWO TYPES OF PAI HAVE BEEN IDENTIFIED. PAI-1 IS AN
 CC ACID-STABLE GLYCOPROTEIN FOUND IN PLASMA AND PLATELETS AND IN
 CC ENDOTHELIAL, HEPATOMA, AND FIBROSARCOMA CELLS.
 CC
 CC -1- MISCELLANEOUS: VASCULAR ENDOTHELIAL CELLS MAY BE THE PRIMARY SITE
 CC OF SYNTHESIS OF PLASMA PAI.
 CC
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.lsb.slb.ch/announce/>
 CC or send an email to license@lsb.slb.ch).
 CC
 CC EMBL; X04429; CAA28025.1; -
 CC EMBL; M14083; AAB60008.1; -
 CC EMBL; X04729; CAA28438.1; -
 CC EMBL; X04731; CAA28442.1; -
 CC EMBL; M15006; AAB60003.1; -
 CC EMBL; M22314; AAB60009.1; -
 CC EMBL; M22315; AAB60009.1; JOINED.
 CC EMBL; M22316; AAB60009.1; JOINED.
 CC EMBL; M22317; AAB60009.1; JOINED.
 CC EMBL; M22318; AAB60009.1; JOINED.
 CC EMBL; M22319; AAB60009.1; JOINED.
 CC EMBL; M22320; AAB60009.1; JOINED.
 CC EMBL; X13323; -; NOT ANNOTATED_CDS.
 CC EMBL; X13338; CAA31732.1; -
 CC EMBL; X13339; CABS1659.1; -
 CC EMBL; X13340; CABS1737.1; -
 CC EMBL; X13341; CABS1606.1; -
 CC EMBL; X13342; CABS1607.1; -
 CC EMBL; X13343; CABS1739.1; -
 CC EMBL; X13344; CABS1739.1; -
 CC EMBL; X13345; CAA31729.1; -
 CC EMBL; X12701; CAA31729.1; -
 CC EMBL; X03764; AAB60007.1; -
 CC EMBL; X04744; CAA28444.1; -
 CC EMBL; AC004876; AAD45828.1; -
 CC PIR; A26996; TRHPI1.
 CC PIR; S02551; S02551.
 CC PDB; 9PAI; 19-MAR-99.
 CC PDB; 1ATC; 23-MAR-99.
 CC PDB; 1B3K; 10-DEC-99.
 CC SWISS-2DPAGE; P05121; HUMAN.
 CC MIM; 173360; -
 CC InterPro; IPR000215; Serpin.
 CC Pfam; PF00079; Serpin; 1.
 CC SMART; SM00093; SERPIN; 1.
 CC PROSITE; PS00284; SERPIN; 1.
 CC Serpin, serine protease inhibitor; Plasma: Plasminogen activation;
 CC Glycoprotein; Polymorphism; Signal; 3p-structure.
 CC SIGNAL 1 23
 CC CHAIN 24 402
 CC ACT_SITE 369 370
 CC CARBOHYD 232 232
 CC CARBOHYD 288 288
 CC CARBOHYD 352 352
 CC
 CC PLASMINOGEN ACTIVATOR INHIBITOR-1.
 CC REACTIVE BOND.
 CC N-LINKED (GLCNAc. . .) (POTENTIAL).
 CC N-LINKED (GLCNAc. . .) (POTENTIAL).
 CC N-LINKED (GLCNAc. . .) (POTENTIAL).
 CC N-LINKED (GLCNAc. . .) (POTENTIAL).

FT VARIANT 15 15 A->T.
/FTID=VAR_007099.

Query Match 100.0%; Score 55; DB 1; Length 402;
Best Local Similarity 100.0%; Pred. No. 0.0017;
Matches 11: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RMAPEEIMDR 11
DB 369 RMAPEEIMDR 379

RESULT 4
ID PA1L_PIG STANDARD; PRT: 402 AA.
AC P79335;
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PLASMINOGEN ACTIVATOR INHIBITOR-1 PRECURSOR (PAI-1) (ENDOTHELIAL
SERPINE1 OR PAI1 OR PLANH1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
NCBI_TaxID=9823;
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=97205538; PubMed=9157595;
RA Bjijens A.P., Knockaert I., Cousin E., Krulthof E.K.O., Declercq P.J.;
RT Expression and characterization of recombinant porcine plasminogen
activator inhibitor-1.
RL Thromb. Haemost. 77:350-356(1997).
[2]
RP ERRATUM.
RA Bjijens A.P., Knockaert I., Cousin E., Krulthof E.K.O., Declercq P.J.;
RL Thromb. Haemost. 77:1046-1046(1997).
CC -1- FUNCTION: THIS INHIBITOR ACTS AS "BAIT" FOR TISSUE PLASMINOGEN
ACTIVATOR, UROKINASE, AND PROTEIN C. ITS RAPID INTERACTION WITH
TPA MAY FUNCTION AS A MAJOR CONTROL POINT IN THE REGULATION OF
FIBRINOLYSIS.
CC -1- MISCELLANEOUS: PAI1 IS INACTIVATED BY PROTEOLYTIC ATTACK OF THE
UROKINASE-TYPE (U-PA) AND THE TISSUE-TYPE (TPA), CLEAVING THE
369(R)-370(M) BOND (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL: Y11347; CAA72182.1; -.
DR HSSP: P01008; IATH.
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; serpin.1.
DR SMART: SM00093; SERPIN.1.
DR PROSITE: PS00284; SERPIN.1.
KW Serpin; Serine protease inhibitor; Plasma; Plasminogen activation;
KW Glycoprotein; Signal.
FT SIGNAL 1 23
FT CHAIN 24 402 BY SIMILARITY
FT ACT_SITE 369 370 PLASMINOGEN ACTIVATOR INHIBITOR-1.
FT CARBOHYD 232 232 REACTIVE BOND.
FT CARBOHYD 288 288 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 402 AA; 45450 MW; 13F6055F4F9FE405 CRC64;

Query Match 100.0%; Score 55; DB 1; Length 402;
Best Local Similarity 100.0%; Pred. No. 0.0017;

Matches 11: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RMAPEEIMDR 11
DB 369 RMAPEEIMDR 379

RESULT 5
ID PA1L_RAT STANDARD; PRT: 402 AA.
AC P20961;
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PLASMINOGEN ACTIVATOR INHIBITOR-1 PRECURSOR (PAI-1) (ENDOTHELIAL
SERPINE1 OR PAI1 OR PLANH1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
RN
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90130456; PubMed=2298740;
RA Brudzinski C.J., Riordan-Johnson M., Nordby E.C., Suter S.M.,
RA Gelehrter T.D.;
RT "Isolation and characterization of the rat plasminogen activator
inhibitor-1 gene."
RL J. Biol. Chem. 265:2078-2085(1990).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89211983; PubMed=3149611;
RA Zehet R., Gelehrter T.D.;
RT "Cloning and sequencing of cDNA for the rat plasminogen activator
inhibitor-1."
RL Gene 73:459-468(1988).
CC -1- FUNCTION: THIS INHIBITOR ACTS AS "BAIT" FOR TISSUE PLASMINOGEN
ACTIVATOR, UROKINASE, AND PROTEIN C. ITS RAPID INTERACTION WITH
TPA MAY FUNCTION AS A MAJOR CONTROL POINT IN THE REGULATION OF
FIBRINOLYSIS.
CC -1- MISCELLANEOUS: PAI1 IS INACTIVATED BY PROTEOLYTIC ATTACK OF THE
UROKINASE-TYPE (U-PA) AND THE TISSUE-TYPE (TPA), CLEAVING THE
369(R)-370(M) BOND.
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL: J05206; AAA11796.1; -.
DR EMBL: M2406; AAA5656.1; -.
DR PIR: A35032; A35032.
DR HSSP: P01008; IATH.
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; serpin.1.
DR SMART: SM00093; SERPIN.1.
DR PROSITE: PS00284; SERPIN.1.
KW Serpin; Serine protease inhibitor; Plasma; Plasminogen activation;
KW Glycoprotein; Signal.
FT SIGNAL 1 23
FT CHAIN 24 402 BY SIMILARITY
FT ACT_SITE 369 370 PLASMINOGEN ACTIVATOR INHIBITOR-1.
FT CARBOHYD 232 232 REACTIVE BOND.
FT CARBOHYD 288 288 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 402 AA; 45009 MW; 2B639140F475EB53 CRC64;

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: M33960; AAA39887.1; -
 DR PIR: A34761; A34761.
 DR HSP: P01008; IATH.
 DR MGI: MGI:97608; Serpinel.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin.1.
 DR SMART: SM00093; SERPIN.1.
 DR PROSITE: PS00284; SERPIN.1.
 DR Serpin: Serpin protease inhibitor; Plasma; plasminogen activation;
 KW Glycoprotein; Signal.
 FT SIGNAL 1 23
 FT CHAIN 1 402
 FT ACT_SITE 369 370 PLASMINOGEN ACTIVATOR INHIBITOR-1.
 FT CARBOHYD 232 232 REACTIVE BOND.
 FT CARBOHYD 288 288 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 402 AA; 45170 MW; 765FF1659C70F68C CRC64;
 SQ SEQUENCE

Query Match 74.5% Score 41; DB 1; Length 402;

Best Local Similarity 63.6% Pred. No. 1.1; Mismatches 0; Gaps 0;

Matches 7; Conservative 3; Indels 1; Indels 0; Gaps 0;

OY 1 MAPEITMDR 11
 |||||:::
 DB 369 RMAPELVDR 379

RESULT 8
 CATA_CAUCR STANDARD; PRT; 737 AA.
 AC 031066;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DE 20-AUG-2001 (Rel. 40, Last annotation update)
 DE PEROXIDASE/CATALASE (EC 1.11.1.6) (CATALASE-PEROXIDASE).
 GN KATG OR CC3043.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
 OC Caulobacter.
 OX NCBI_TaxID=69394;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 19089 / CB15;
 RX MEDLINE-21173698; PubMed-11259647;
 RA Eisen J., Heidelberg J.F., Alley M.K.R., Ohta N., Maddock J.R.,
 RA Pirocha I., Nelson M.C., Newton A., Stephens C., Phade N.D., Ely B.,
 RA Deboy R.T., Dodson R.J., Durkin A.S., Gilm M.L., Haft D.H.,
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shelly J., Berry K.,
 RA Uetebach T., Tran K., Wolf A., Yamashiro J., Ermolaeva M., White O.,
 RA Salzberg S.L., Shapiro L., Venter J.C., Fraser C.M.;
 RA "Complete genome sequence of Caulobacter crescentus."
 RT Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 RL [2]
 RN SEQUENCE OF 1-494 FROM N.A.
 RP STRAIN-ATCC 19089 / CB15;
 RX MEDLINE-98012985; PubMed-9352936;
 RA Steinman H.M., Farred F., Weinstein L.;
 RA "Catalase-peroxidase of Caulobacter crescentus: function and role in
 RT stationary-phase survival."
 RL J. Bacteriol. 179:6831-6836(1997).
 CC -1- FUNCTION: BIFUNCTIONAL, EXHIBITING BOTH A CATALASE AND
 CC BROAD-SPECTRUM PEROXIDASE ACTIVITIES.
 CC -1- CATALYTIC ACTIVITY: DONOR + H(2)O(2) -> OXIDIZED DONOR + 2 H(2)O.
 CC -1- COFACTOR: TWO PROTOHEME IX GROUPS (BY SIMILARITY).
 CC -1- INDUCTION: BY EXPOSURE TO HYDROGEN PEROXIDE.

-1- SIMILARITY: BELONGS TO THE PEROXIDASE FAMILY. BACTERIAL
 PEROXIDASE/CATALASE SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: AE005967; AAK25005.1; -
 DR EMBL: AF027168; AAC45850.1; -
 DR HSP: P48534; IAPX.
 DR TIGR: CC3043; -
 DR InterPro: IPR002016; Peroxidase.
 DR Pfam: PF00141; peroxidase.1.
 DR PROSITE: PS00435; PEROXIDASE_1; 1.
 DR PROSITE: PS00436; PEROXIDASE_2; 1.
 KW Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
 FT ACT_SITE 102 102
 FT ACT_SITE 106 106 BY SIMILARITY.
 FT BINDING 268 268 BY SIMILARITY.
 FT BINDING 268 268 PROXIMAL HEME LIGAND (BY SIMILARITY).
 SQ SEQUENCE 737 AA; 80165 MW; D900742D04E02E97 CRC64;

Query Match 70.9% Score 39; DB 1; Length 737;

Best Local Similarity 70.0% Pred. No. 5.4; Mismatches 1; Indels 0; Gaps 0;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 MAPEITMDR 11
 |||||:::
 DB 601 MAPEALVDR 610

RESULT 9
 YPH5_THIYI STANDARD; PRT; 692 AA.
 ID YPH5_THIYI
 AC P45365;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE HYPOETHERICAL 76.5 KDA PROTEIN IN PHBC 3' REGION (ORF5).
 OS Thlocystis violacea.
 OC Bacteria; Proteobacteria; gamma subdivision; Chromatiaceae;
 OC Thlocystis.
 OX NCBI_TaxID=13725;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-2311 / DSM 208;
 RX MEDLINE-93159750; PubMed-7763384;
 RA Liebergesell M., Steinbuechel A.;
 RA "Cloning and molecular analysis of the poly(3-hydroxybutyric acid)
 RT biosynthetic genes of Thlocystis violacea."
 RL Appl. Microbiol. Biotechnol. 38:493-501(1993).
 CC -1- SIMILARITY: THE N-TERMINAL REGION IS SIMILAR TO THAT OF OTHER
 CC REGULATORY COMPONENTS OF SENSOR TRANSDUCTION SYSTEMS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: L01113; AAB02863.1; -
 DR EMBL: S54369; AAC60451.1; -
 DR InterPro: IPR001633; DUF2.
 DR InterPro: IPR001633; DUF2.
 DR InterPro: IPR001789; Response_reg.
 DR Pfam: PF00563; DUF2; 1.

Best Local Similarity 50.0%; Pred. No. 43;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 2 MAPEIIMDR 11
|:|:|:|:|:
DB 499 MSPEITLMDK 508

RESULT 12

RRPL_TOSV STANDARD; PRT: 2095 AA.
ID RRPL_TOSV
AC P37800;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.48) (L PROTEIN).
GN L. Toscana virus (Tos).
OS Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.
OX NCBI_TaxID=11590;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISS.PHL.3;
RA MEDLINE=9312587; PubMed=8460526;
RA Accardi L., Gro M.C., di Bonito P., Giorgi C.;
RT "Toscana virus genomic L segment: molecular cloning, coding strategy
RT and amino acid sequence in comparison with other negative strand RNA
RT viruses."; 27:119-131(1993)
RL Virus Res.
CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE - N PROPHOSPHATE
CC + RNA(N).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X68414; CAA6478.1; -
DR PIR: S29529; S29529.
KW Transferase: Nucleotidyltransferase; RNA-directed RNA polymerase.
SQ SEQUENCE 2095 AA; 23885 MW; 22FF4DAD745583B7 CRC64;

Query Match 63.6%; Score 35; DB 1; Length 2095;
Best Local Similarity 45.5%; Pred. No. 1; Indels 0; Gaps 0;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 1 RMAPEIIMDR 11
|:|:|:|:|:
DB 1519 RVAPETLSDK 1529

RESULT 13

MOBB_ECOLI STANDARD; PRT: 174 AA.
ID MOBB_ECOLI
AC P31125; P76770;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE MOLYDOPTERIN-GUANINE DINUCLEOTIDE BIOSYNTHESIS PROTEIN B.
GN MOBB OR B3856.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RA MEDLINE=93347969; PubMed=8346018;
RA Plunkett G. III, Burland V.D., Daniels D.L., Blattner F.R.;

"Analysis of the Escherichia coli genome. III. DNA sequence of the
RT region from 87.2 to 89.2 minutes."
RL Nucleic Acids Res. 21:3391-3398(1993).

RP SEQUENCE FROM N.A., SEQUENCE OF 1-7, AND CHARACTERIZATION.
RN [2]

RA MEDLINE=96004466; PubMed=7551035;
RA Lohbi-Nivol C., Palmer T., Whitty P.W., McNaughton E., Boxer D.H.;
RT "The mob locus of Escherichia coli K12 required for molybdenum
RT cofactor biosynthesis is expressed at very low levels."
RL Microbiology 141:1663-1671(1995).

RM CHARACTERIZATION.

RA MEDLINE=97363202; PubMed=9219527;
RA Eaves D.J., Palmer T., Boxer D.H.;

RT "The product of the molybdenum cofactor gene mobB of Escherichia coli
RT is a GTP-binding protein."
RL Eur. J. Biochem. 246:690-697(1997).

CC -1- FUNCTION: MAY BIND THE GUANINE NUCLEOTIDE REQUIRED FOR THE
CC SYNTHESIS OF MOLYDOPTERIN GUANINE DINUCLEOTIDE.

CC -1- PATHWAY: MOLYBDENUM COFACTOR BIOSYNTHESIS.

CC -1- SUBUNIT: HOMODIMER.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL: L19201; AAB02991.1; ALT_INIT.

DR EMBL: AE000461; AAC76854.1; ALT_INIT.

DR PIR: S40802; S40802.

DR EcGene: EG1828; mobB.

KW Molybdenum cofactor biosynthesis; GTP-binding; Complete proteome.

FT INT_MET 0

FT NP_BIND 13 20 GTP (POTENTIAL).

FT NP_BIND 50 53 GTP (POTENTIAL).

FT NP_BIND 99 102 GTP (POTENTIAL).

FT NP_BIND 174 AA; 19232 MW; C3522F4DFB3833C CRC64;

Query Match 61.8%; Score 34; DB 1; Length 174;
Best Local Similarity 62.5%; Pred. No. 11;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 4 PEETIMDR 11
|:|:|:|:|:
DB 128 PEELVIDR 135

RESULT 14

RANT_MOUSE STANDARD; PRT: 216 AA.
ID RANT_MOUSE
AC O61820;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GTP-BINDING NUCLEAR PROTEIN RAN, TESTIS-SPECIFIC ISOFORM.
GN RANL2-9.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA MEDLINE=95152168; PubMed=7849398;
RA Coutavas E.E., Hsien C.M., Ren M., Drivas G.T., Rush M.G.,
RA D'Eustachio P.D.;
RT "Tissue-specific expression of Ran isoforms in the mouse."
RL Mamm. Genome 5:623-628(1994).
CC -1- FUNCTION: GTP-BINDING PROTEIN INVOLVED IN NUCLEOCYTOPLASMIC

CC TRANSFORM. REQUIRED FOR THE IMPORT OF PROTEIN INTO THE NUCLEUS AND
 CC ALSO FOR RNA EXPORT. INVOLVED IN CHROMATIN CONDENSATION AND
 CC CONTROL OF CELL CYCLE (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- TISSUE SPECIFICITY: TESTIS-SPECIFIC.
 CC -1- SIMILARITY: BELONGS TO THE RAN FAMILY IN THE RAS SUPERFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: L32752; AAA64248.1; .
 CC HSSP: P28746; 1A2K.
 CC MGD: MGI:104605; Ras12-9.
 CC InterPro: IPR002041; RAN.
 CC InterPro: IPR001806; Ras_trnsfmg.
 CC InterPro: IPR003575; Small_GTPase.
 CC Pfam: PF00071; Ras; 1.
 CC PRINTS: PR00449; RASTRNSFRNG.
 CC PRINTS: PR00627; GTPRANTC4.
 CC SMART: SM00176; RAN; 1.
 CC SMART: SM00010; small_GTPase; 1.
 CC PROSITE: PS0115; RAN; 1.
 CC GTP-binding; Nuclear protein; Protein transport.
 CC NP_BIND 17 24 GTP (BY SIMILARITY).
 CC NP_BIND 65 69 GTP (BY SIMILARITY).
 CC NP_BIND 122 125 GTP (BY SIMILARITY).
 CC DOMAIN 127 142 IBB (BY SIMILARITY).
 CC SEQUENCE 216 AA; 24452 MW; 23BA7E723B96B808 CRC64;
 SO
 Query Match 61.8%; Score 34; DB 1; Length 216;
 Best Local Similarity 55.6%; Pred. No. 14;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 Oy 2 MAPEIIMD 10
 Db 182 LAPPEVMD 190
 RESULT 15
 RAN_CHICK ID RAN_CHICK STANDARD; PRT; 216 AA.
 AC P42558;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE GTP-BINDING NUCLEAR PROTEIN RAN (TC4).
 GN RAN.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Archosaustra; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 CC NCBI_TaxID=9031;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC MEDLINE=92339546; PubMed=1633874;
 CC Truub J., Truub B.;
 CC FESB Lett. 306:181-184(1992).
 CC -1- FUNCTION: GTP-BINDING PROTEIN INVOLVED IN NUCLEOCYTOPLASMIC
 CC TRANSPORT. REQUIRED FOR THE IMPORT OF PROTEIN INTO THE NUCLEUS AND
 CC ALSO FOR RNA EXPORT. INVOLVED IN CHROMATIN CONDENSATION AND
 CC CONTROL OF CELL CYCLE (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- SIMILARITY: BELONGS TO THE RAN FAMILY IN THE RAS SUPERFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X65906; CAA47355.1; .
 CC HSSP: P28746; 1A2K.
 CC InterPro: IPR002041; RAN.
 CC InterPro: IPR001806; Ras_trnsfmg.
 CC Pfam: PF00071; Ras; 1.
 CC PRINTS: PR00449; RASTRNSFRNG.
 CC PRINTS: PR00627; GTPRANTC4.
 CC SMART: SM00176; RAN; 1.
 CC PROSITE: PS0115; RAN; 1.
 CC GTP-binding; Nuclear protein; Protein transport.
 CC NP_BIND 17 24 GTP (BY SIMILARITY).
 CC NP_BIND 65 69 GTP (BY SIMILARITY).
 CC NP_BIND 122 125 GTP (BY SIMILARITY).
 CC DOMAIN 127 142 IBB (BY SIMILARITY).
 CC SEQUENCE 216 AA; 24427 MW; D5DB00D75C2C0DE4 CRC64;
 SO

Query Match 61.8%; Score 34; DB 1; Length 216;
 Best Local Similarity 55.6%; Pred. No. 14;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 Oy 2 MAPEIIMD 10
 Db 182 LAPPEVMD 190

Search completed: March 18, 2002, 09:52:25
 Job time: 294 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 18, 2002, 09:51:54 ; Search time 78.82 Seconds
(without alignments)
20.414 Million cell updates/sec

Title: US-09-544-665-2
Perfect score: 55
Sequence: 1 RMAPEIIVDR 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL-17:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	37	67.3	382	10	Q9SGC8	Q9SGC8 arabidopsis
2	36	65.5	160	2	Q9GCI6	Q9GCI6 mycobacteri
3	36	65.5	329	5	Q9U1B4	Q9U1B4 leishmania
4	36	65.5	533	10	Q9LKK7	Q9LKK7 arabidopsis
5	36	65.5	534	10	Q9LDH0	Q9LDH0 arabidopsis
6	36	65.5	723	2	Q9R708	Q9R708 agrobacteri
7	35	63.6	220	12	Q990I6	Q990I6 parametium
8	35	63.6	220	12	Q990I8	Q990I8 parametium
9	35	63.6	220	12	Q990I7	Q990I7 parametium
10	35	63.6	220	12	Q990I3	Q990I3 parametium
11	35	63.6	540	1	Q30043	Q30043 archaeoglob
12	35	63.6	624	12	Q9SGS5	Q9SGS5 chilo. iride
13	35	63.6	651	2	Q9F0L8	Q9F0L8 staphylococ
14	35	63.6	1337	4	Q9P2E4	Q9P2E4 homo sapien
15	35	63.6	1499	2	Q99WD1	Q99WD1 staphylococ
16	34	61.8	48	6	Q9TS94	Q9TS94 erythrocyt
17	34	61.8	120	2	Q9A256	Q9A256 caulobacter
18	34	61.8	131	4	Q9UDJ5	Q9UDJ5 homo sapien
19	34	61.8	170	2	Q9LBY1	Q9LBY1 escherichia

20	34	61.8	176	2	Q67460	Q67460 aquifex aeo
21	34	61.8	215	13	Q9YGC0	Q9YGC0 salmo salar
22	34	61.8	216	4	Q9UEU9	Q9UEU9 homo sapien
23	34	61.8	216	11	Q9D9Y0	Q9D9Y0 mus muscu
24	34	61.8	216	11	Q9CWI7	Q9CWI7 mus muscu
25	34	61.8	216	13	Q9IBEB	Q9IBEB xenopus lae
26	34	61.8	250	11	Q9CSP3	Q9CSP3 mus muscu
27	34	61.8	284	10	Q82397	Q82397 arabidopsis
28	34	61.8	285	5	Q94281	Q94281 caenorhabdi
29	34	61.8	348	2	Q9X026	Q9X026 thermotoga
30	34	61.8	398	2	Q9KPE6	Q9KPE6 vibrio chol
31	34	61.8	402	2	Q86587	Q86587 streptomyce
32	34	61.8	402	2	Q88124	Q88124 streptomyce
33	34	61.8	432	5	Q9GRF9	Q9GRF9 leishmania
34	34	61.8	658	8	Q9GDX7	Q9GDX7 erythronum
35	34	61.8	720	2	Q9R6S9	Q9R6S9 synecchococ
36	34	61.8	765	10	Q9FCD7	Q9FCD7 arabidopsis
37	34	61.8	772	1	Q51963	Q51963 halobacteri
38	33	60.0	150	2	Q99RNE	Q99RNE staphylococ
39	33	60.0	168	2	Q9A565	Q9A565 caulobacter
40	33	60.0	215	12	Q990K7	Q990K7 parametium
41	33	60.0	215	12	Q990K6	Q990K6 parametium
42	33	60.0	215	12	Q990K5	Q990K5 parametium
43	33	60.0	215	12	Q990K4	Q990K4 parametium
44	33	60.0	215	12	Q990K3	Q990K3 parametium
45	33	60.0	216	2	Q06392	Q06392 mycobacteri

ALIGNMENTS

RESULT 1
ID Q9SGC8 PRELIMINARY; PRT; 382 AA.
AC Q9SGC8:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE T23G18.18.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI-TaxID=3702;
RX [1]
RN SEQUENCE FROM N.A.
RA Shin P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
RA Kim C., Altafi H., Bel Q., Chin C., Chlou J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Howling B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thayerl A.,
RA Tortumli M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.,
RT "Genomic sequence for Arabidopsis thaliana BAC T23G18 from chromosome
RT 1."
RT Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC011438; AAF18244.1;
DR InterPro; IPR000379; Est_11P_thioest_acsite.
SQ SEQUENCE 382 AA; 43223 MM; 353040AF91B84E0 CRC64;

Query Match 67.3% Score 37; DB 10; Length 382;
Best Local Similarity 70.0%; Pred. No. 60;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 2 RMAPEIIVDR 11
DB 345 IAPREIIVDR 354
RESULT 2
Q9GCI6 PRELIMINARY; PRT; 160 AA.

AC Q9CIC6; 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE HYPOTHETICAL PROTEIN ML0781.
 GN ML0781.
 OS Mycobacterium leprae.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacterium.
 NC NCBI_TaxID=1769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TN;
 RX MEDLINE=21128732; PubMed=11234002;
 RA Cole S.T., Eiglmeyer K., Parkhill J., James K.D., Thomson N.R.,
 RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
 RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
 RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
 RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skellon J., Squares R.,
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
 RA Barrall B.G.;
 RT "Massive gene decay in the leprosy bacillus";
 RL Nature 409:1007-1011(2001).
 DR EMBL: AL583919; CAC30290.1;
 DR InterPro: IPR000309; TrKA_Kruptake.
 DR Pfam: PF02080; TrKA_1.
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 160 AA; 17406 MW; AD845F5F2F616AC CRC64;

Query Match 65.5%; Score 36; DB 2; Length 160;
 Best Local Similarity 54.5%; Pred. No. 37;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

RESULT 3
 ID Q901B4 PRELIMINARY; PRT; 329 AA.
 AC Q901B4;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE POSSIBLE BETA PROPELLER PROTEIN (FRAGMENT).
 GN L2581.14.
 OS Leishmania major.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 NC NCBI_TaxID=5664;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FRIEDLIN;
 RX Zimmermann W., Wambutt R., Ivens A.C., Lawson D., Quail M.,
 RA Rajandream M.A., Barrall B.G.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FRIEDLIN;
 RX MEDLINE=9814635; PubMed=9477341;
 RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
 RA Smith D.F.;
 RT "A physical map of the Leishmania major Friedlin genome";
 RL Genome Res. 8:135-143(1998).
 DR EMBL: AL121851; CAB58300.1;
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00400; WD40; 3.
 DR SMART: SM00320; WD40; 4.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 KM Repeat; WD repeat.

FT NON TER 1 1
 SQ SEQUENCE 329 AA; 36924 MW; 07759E2DFD138FF CRC64;

Query Match 65.5%; Score 36; DB 5; Length 329;
 Best Local Similarity 55.6%; Pred. No. 79;
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 3 APEEIIIMDR 11
 Db 178 APEELVLDK 186

RESULT 4
 ID Q9LKK7 PRELIMINARY; PRT; 533 AA.
 AC Q9LKK7;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE BETA 1,2-XYLOSYLTRANSFERASE.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 NC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV; WS;
 RA Gomord V., Balthresa V., Kiefer-Meyer M.-C., Paye L.;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF272852; AAF77064.1;
 KM Transferase.
 SQ SEQUENCE 533 AA; 60136 MW; 82F617B871FA2DFD CRC64;

Query Match 65.5%; Score 36; DB 10; Length 533;
 Best Local Similarity 54.5%; Pred. No. 1,3e+02;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RMAPEEIIIMDR 11
 Db 127 RMVDPRIYMSR 137

RESULT 5
 ID Q9LDH0 PRELIMINARY; PRT; 534 AA.
 AC Q9LDH0;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE BETA1,2-XYLOSYLTRANSFERASE (GENOMIC DNA, CHROMOSOME 5, P1
 CLONE:MTE17).
 GN XLT.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 NC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20245450; PubMed=10781814;
 RA Strasser R., Mucha J., Mach L., Altman F., Wilson I.B.H., Gloessel J.,
 RA Steinkellner H.;
 RT "Molecular cloning and functional expression of beta 1,2-
 xylosyltransferase cDNA from Arabidopsis thaliana";
 RL FEBS Lett. 472:105-108(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Bakker H.;
 RT "Arabidopsis cDNA for beta1,2-xylosyltransferase";
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.

RN [3]
 RC SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA;
 RX MEDLINE-99087489; PubMed-9872454;
 RA Nakamura Y., Sato S., Asamizu E., Kaneko T., Kotani H., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. VII.
 RT Sequence features of the regions of 1,013,767 bp covered by sixteen
 RT physically assigned P1 and TAC clones.";
 RL DNA Res. 5:297-308(1998).
 DR EMBL: AJ272121; CAB90610.1; -;
 DR EMBL: AJ277603; CAB9489.1; -;
 DR EMBL: AB015479; BAB08567.1; -;
 KW TRANSFERASE
 SQ SEQUENCE 534 AA; 60234 MW; A7DCA7FDE8365874 CRC64;

Query Match 65.5%; Score 36; DB 10; Length 534;
 Best Local Similarity 54.5%; Pred. No. 1,3e+02;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 MAFPEIIMDR 11
 DB 127 RMVDPRIWMSR 137

RESULT 6
 ID 098708 PRELIMINARY; PRT: 723 AA.
 AC 098708;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE CATALASE.
 GN KATA.
 OS Acetabacterium radiobacter.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=58;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Xu X.O., Pan S.O.;
 RT "An Agrobacterium catalase is a virulence factor involved in
 RT tumorigenesis.";
 RL Mol. Microbiol. 0:0-0(1999).
 DR EMBL: AB033631; BAA89349.1; -;
 DR HSP: P48534; IAPY.
 DR InterPro: IPR002016; Peroxidase.
 DR InterPro: IPR001680; MD40.
 DR Pfam: PF00141; peroxidase; 1.
 DR PRINTS: PRO0436; PEROXIDASE.
 DR PROSITE: PS00436; PEROXIDASE_2; UNKNOWN_1.
 DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_1.
 SQ SEQUENCE 723 AA; 78741 MW; AA51987F53B508 CRC64;

Query Match 65.5%; Score 36; DB 2; Length 723;
 Best Local Similarity 50.0%; Pred. No. 1.8e+02;
 Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

OY 2 MAFPEIIMDR 11
 DB 588 VSPPELIDR 597

RESULT 7
 ID 099016 PRELIMINARY; PRT: 220 AA.
 AC 099016;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE DNA POLYMERASE (FRAGMENT);

GN DPO.
 OS Parametium buraria chlorella virus 1 (PCV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Phycodnavirus.
 OX NCBI_TaxID=10506;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AR158;
 RA Zhang Y., Adams B., Sun L., Burbank D., Van Etten J.L.;
 RT "Intron conservation in the DNA polymerase gene encoded by Chlorella
 RT viruses.";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF344202; AAK28923.1; -;
 FT NON-TER 1
 FT NON-TER 220
 SQ SEQUENCE 220 AA; 24907 MW; 6FF0FD7EC4C50D9 CRC64;

Query Match 63.6%; Score 35; DB 12; Length 220;
 Best Local Similarity 50.0%; Pred. No. 80;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 2 MAFPEIIMDR 11
 DB 27 MSPETLVMDK 36

RESULT 8
 ID 099018 PRELIMINARY; PRT: 220 AA.
 AC 099018;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE DNA POLYMERASE (FRAGMENT).
 GN DPO.
 OS Parametium buraria chlorella virus 1 (PCV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Phycodnavirus.
 OX NCBI_TaxID=10506;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NT-2A;
 RA Zhang Y., Adams B., Sun L., Burbank D., Van Etten J.L.;
 RT "Intron conservation in the DNA polymerase gene encoded by Chlorella
 RT viruses.";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF344230; AAK28951.1; -;
 FT NON-TER 1
 FT NON-TER 220
 SQ SEQUENCE 220 AA; 24945 MW; 691DB6C65E39B7B5 CRC64;

Query Match 63.6%; Score 35; DB 12; Length 220;
 Best Local Similarity 50.0%; Pred. No. 80;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 2 MAFPEIIMDR 11
 DB 27 MSPETLVMDK 36

RESULT 9
 ID 099017 PRELIMINARY; PRT: 220 AA.
 AC 099017;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE DNA POLYMERASE (FRAGMENT).
 GN DPO.
 OS Parametium buraria chlorella virus 1 (PCV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Phycodnavirus.
 OX NCBI_TaxID=10506;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-NY-2B;
 RT Zhang Y., Adams B., Sun L., Burbank D., Van Etten J.L.;
 RT "Inton conservation in the DNA polymerase gene encoded by Chlorella
 RT viruses.";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF344231; AAK26952.1;
 FT NON_TER 1
 FT SEQUENCE 220 AA; 24985 MW; 73C8C7AB3B4056F CRC64;

Query Match 63.6%; Score 35; DB 12; Length 220;
 Best Local Similarity 50.0%; Pred. No. 80;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 2 MAPEIIMDR 11
 DB 27 MSPETLVMK 36

RESULT 10
 ID 099013 PRELIMINARY; PRT; 220 AA.
 AC 099013;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE DNA POLYMERASE (FRAGMENT).
 OS DPO.
 OS Parametrium bursaria chlorella virus 1 (PBCV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Phycodnavirus.
 OX NCBI_TaxID=10506;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NYS-1;
 RT Zhang Y., Adams B., Sun L., Burbank D., Van Etten J.L.;
 RT "Inton conservation in the DNA polymerase gene encoded by Chlorella
 RT viruses.";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF344235; AAK26956.1;
 FT NON_TER 1
 FT SEQUENCE 220 AA; 24927 MW; 7135E9B62C54366F CRC64;

Query Match 63.6%; Score 35; DB 12; Length 220;
 Best Local Similarity 50.0%; Pred. No. 80;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 2 MAPEIIMDR 11
 DB 27 MSPETLVMK 36

RESULT 11
 ID 030043 PRELIMINARY; PRT; 540 AA.
 AC 030043;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE MEDICUM-CHAIN ACYL-COA LIGASE (ALKK-1).
 GN AF0196.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
 OC Archaeoglobus
 OX NCBI_TaxID=2234;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,

RA Ketchum K.A., Dodson R.J., Gwin M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kierlage A.R., Graham D.E., Kyriades N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.O., Gill S.,
 RA Kirschner E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
 RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaeon Archaeoglobus fulgidus.";
 RL Nature 390:364-370(1997).
 DR EMBL: AEO01093; AAB91037.1;
 DR HSSP: P08659; 1ICI.
 DR TIGR: AF0196;
 DR InterPro: IPR000873; AMP-bind.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP_BINDING; 1.
 KW Hypothetical protein; Ligase; Complete proteome.
 SQ SEQUENCE 540 AA; 61888 MW; 32A4962CA45E6466 CRC64;

Query Match 63.6%; Score 35; DB 1; Length 540;
 Best Local Similarity 75.0%; Pred. No. 2,1e+02;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 RMAPEIIV 8
 DB 87 RLAPEIIV 94

RESULT 12
 ID 0908K5 PRELIMINARY; PRT; 624 AA.
 AC 0908K5;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE EXONUCLEASE II HOMOLOG.
 GN EXO2.
 OS Chilo iridescent virus (CIV) (Insect iridescent virus type 6).
 OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.
 OX NCBI_TaxID=10488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Muller K., Tidona C.A., Bahr U., Darai G.;
 RT "The DNA sequence of Chilo iridescent virus between the genome
 RT coordinates 0.974 and 0.101.";
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF083915; AAD48147.1;
 KW Exonuclease.
 SQ SEQUENCE 624 AA; 72576 MW; 64EB22157D5A5D9A CRC64;

Query Match 63.6%; Score 35; DB 12; Length 624;
 Best Local Similarity 66.7%; Pred. No. 2,4e+02;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 3 APEIIMDR 11
 DB 559 APEIIMDR 567

RESULT 13
 ID 09F0L8 PRELIMINARY; PRT; 691 AA.
 AC 09F0L8;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE GLUTAMATE SYNTHASE LARGE SUBUNIT (FRAGMENT).
 GN GLTB.

OS Staphylococcus aureus.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Staphylococcus.
 OX NCBI_TaxID=1280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Boyle-Vavra S., Ebert C.E., Daum R.S.;
 RT "The gltB gene encoding the large (a) subunit of glutamate synthase
 identified by screening a vancomycin-resistant Staphylococcus aureus
 library on medium containing vancomycin."
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF217184; AAC44102.1;
 FT NON_TER 1
 FT SEQUENCE 691 AA; 77731 MW; DC7BCE2804890C18 CRC64;

Query Match 63.6%; Score 35; DB 2; Length 691;
 Best Local Similarity 50.0%; Pred. No. 2.7e+02;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 2 MAPEIIMDR 11
 DB 494 LAPDETVDLR 503

RESULT 14
 ID Q9P2E4 PRELIMINARY; PRT; 1337 AA.
 AC Q9P2E4;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE KIAA1403 PROTEIN (FRAGMENT).
 GN KIAA1403.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RX MEDLINE=20181126; PubMed=10718198;
 RA Nagase T., Kikuno R., Ishikawa K., Hirose M., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XVI.
 The complete sequences of 150 new cDNA clones from brain which code
 for large proteins in vitro."
 RL DNA Res. 7:65-73(2000).
 DR EMBL: AB037824; BAA92641.1;
 FT NON_TER 1
 FT SEQUENCE 1337 AA; 146216 MW; 17DF1DE835E06E01 CRC64;

Query Match 63.6%; Score 35; DB 4; Length 1337;
 Best Local Similarity 60.0%; Pred. No. 5.5e+02;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 MAPEIIMDR 11
 DB 260 MAPEIILQEQ 269

RESULT 15
 ID Q99WD1 PRELIMINARY; PRT; 1499 AA.
 AC Q99WD1;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE GLUTAMATE SYNTHASE LARGE SUBUNIT.
 GN GltB OR SA0430.
 OS Staphylococcus aureus subsp. aureus N315.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;

OC Bacillus/Staphylococcus group; Staphylococcus.
 OX NCBI_TaxID=158879;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
 RA Cui L., Oguchi A., Aoki K.I., Nagai Y., Iino T., Kanamori M.,
 RA Matsumaru H., Maruyama A., Murakami H., Hosoyama A., Mizutani-Uchiyama T.,
 RA Takahashi N.K., Sawano T., Inoue R.I., Kato C., Sekimizu K.,
 RA Hirakawa H., Kuhara S., Goto S., Tabuchi J., Kanehisa M.,
 RA Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M.,
 RA Ogatawara N., Hayashi H., Hiramatsu K.;
 RT "Whole genome sequencing of methicillin-resistant Staphylococcus
 aureus."
 RL Lancet 357:1225-1240(2001).
 DR EMBL: AP003130; BAB41660.1;
 KW Complete proteome.
 SO SEQUENCE 1499 AA; 166259 MW; F20B8287BCE001C CRC64;

Query Match 63.6%; Score 35; DB 2; Length 1499;
 Best Local Similarity 50.0%; Pred. No. 6.2e+02;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 2 MAPEIIMDR 11
 DB 541 LAPDETVDLR 550

Search completed: March 18, 2002, 09:51:56
 Job time: 285 sec

Tue Mar 19 11:00:03 2002

us-09-544-665-2.rpt

Page 6

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:48:55 ; Search time 81.32 Seconds
(without alignments)
5.465 Million cell updates/sec

Title: US-09-544-665-3
Perfect score: 29
Sequence: 1 ERIWD 6

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 522463 seqs, 74073290 residues
Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SIDSB/gcgdata/geneseq/AA1980.DAT.*
2: /SIDSB/gcgdata/geneseq/AA1981.DAT.*
3: /SIDSB/gcgdata/geneseq/AA1982.DAT.*
4: /SIDSB/gcgdata/geneseq/AA1983.DAT.*
5: /SIDSB/gcgdata/geneseq/AA1984.DAT.*
6: /SIDSB/gcgdata/geneseq/AA1985.DAT.*
7: /SIDSB/gcgdata/geneseq/AA1986.DAT.*
8: /SIDSB/gcgdata/geneseq/AA1987.DAT.*
9: /SIDSB/gcgdata/geneseq/AA1988.DAT.*
10: /SIDSB/gcgdata/geneseq/AA1989.DAT.*
11: /SIDSB/gcgdata/geneseq/AA1990.DAT.*
12: /SIDSB/gcgdata/geneseq/AA1991.DAT.*
13: /SIDSB/gcgdata/geneseq/AA1992.DAT.*
14: /SIDSB/gcgdata/geneseq/AA1993.DAT.*
15: /SIDSB/gcgdata/geneseq/AA1994.DAT.*
16: /SIDSB/gcgdata/geneseq/AA1995.DAT.*
17: /SIDSB/gcgdata/geneseq/AA1996.DAT.*
18: /SIDSB/gcgdata/geneseq/AA1997.DAT.*
19: /SIDSB/gcgdata/geneseq/AA1998.DAT.*
20: /SIDSB/gcgdata/geneseq/AA1999.DAT.*
21: /SIDSB/gcgdata/geneseq/AA2000.DAT.*
22: /SIDSB/gcgdata/geneseq/AA2001.DAT.*

Printed. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	100.0	5	20	AAV01812
2	29	100.0	33	22	AAV02738
3	29	100.0	33	22	AAV02200
4	29	100.0	220	21	AAV02041
5	29	100.0	236	21	AAV02040
6	29	100.0	284	21	AAV02039
7	29	100.0	379	12	AAV01756
8	29	100.0	379	12	AAV0162
9	29	100.0	379	12	AAV0163
10	29	100.0	379	18	AAV02710
11	29	100.0	379	18	AAV02711

12	29	100.0	379	18	AAV02712	Plasminogen activa
13	29	100.0	379	18	AAV02713	Plasminogen activa
14	29	100.0	379	18	AAV02714	Plasminogen activa
15	29	100.0	379	18	AAV02715	Plasminogen activa
16	29	100.0	379	18	AAV02716	Plasminogen activa
17	29	100.0	379	18	AAV02717	Plasminogen activa
18	29	100.0	379	18	AAV02718	Plasminogen activa
19	29	100.0	379	18	AAV02719	Plasminogen activa
20	29	100.0	379	18	AAV02720	Plasminogen activa
21	29	100.0	379	18	AAV02721	Plasminogen activa
22	29	100.0	379	22	AAV04918	Human Plasminogen
23	29	100.0	379	22	AAV04919	Human Plasminogen
24	29	100.0	379	22	AAV04920	Human Plasminogen
25	29	100.0	379	22	AAV04921	Human Plasminogen
26	29	100.0	379	22	AAV04922	Human Plasminogen
27	29	100.0	379	22	AAV04923	Human Plasminogen
28	29	100.0	379	22	AAV04924	Human Plasminogen
29	29	100.0	379	22	AAV04925	Human Plasminogen
30	29	100.0	379	22	AAV04926	Human Plasminogen
31	29	100.0	379	22	AAV04927	Human Plasminogen
32	29	100.0	379	22	AAV04928	Human Plasminogen
33	29	100.0	379	22	AAV04929	Human Plasminogen
34	29	100.0	382	11	AAV08411	Modified mature pI
35	29	100.0	383	9	AAV08207	Beta plasminogen a
36	29	100.0	402	9	AAV08179	Sequence of human
37	29	100.0	402	11	AAV07986	Plasminogen activa
38	29	100.0	402	18	AAV03587	Human plasminogen
39	29	100.0	402	22	AAV04913	Human plasminogen
40	29	100.0	402	22	AAV07994	Human shear stress
41	27	93.1	319	21	AAV04613	Arabidopsis thalia
42	27	93.1	320	21	AAV037027	Arabidopsis thalia
43	26	89.7	66	21	AAV05541	S. cerevisiae cyto
44	26	89.7	120	21	AAV01409	S. cerevisiae cyto
45	26	89.7	163	22	AAV06542	Human GTP-binding

ALIGNMENTS

RESULT 1	
AAV01812	standard; peptide; 6 AA.
ID	
AAV01812	
AC	AAV01812
XX	
AC	AAV01812
XX	
DT	29-JUN-1999 (first entry)
XX	
DE	Peptide used to regulate scupa.
XX	
KW	Abnormal cell migration; PAI-1-dependent cell adhesion; scupa clearance;
KW	pathological cell migration; angiogenesis; organogenesis; ovulation;
KW	Inflammation; cancer; tumor cell invasion; metastasis; atherosclerosis.
XX	
OS	Synthetic.
XX	
PN	W09920295-A1.
XX	
PD	29-APR-1999.
XX	
PF	15-OCT-1998; 98WO-US21800.
XX	
PR	17-OCT-1997; 97US-0062274.
XX	
PA	(TYPE-) UNTV PENNSYLVANIA.
XX	
PI	Claes D, Higezi AA:
XX	
DR	WPI; 1999-288168/24.
XX	
PT	Peptide composition
XX	
PS	Example 1; Page 31; 63pp; English.
XX	

CC Peptides AAY01812-14 exemplify the peptide of the invention, which is
 CC used in a peptide composition to affect a biological process
 CC characterized by abnormal cell migration through a physiological barrier,
 CC to inhibit PAR-1-dependent adhesion of a cell to a tissue, to promote
 CC clearance of scapA from the surface of a mammalian cell, and to impede
 CC paracrine migration of a cell in a mammal. The composition can be used
 CC to affect a biological process characterized by abnormal cell migration
 CC through a physiological barrier, where the process is selected from cell
 CC angiogenesis, organogenesis, ovulation, inflammation, cancer, tumor cell
 CC invasion and metastasis, and atherosclerosis.

XX Sequence 6 AA;
 SQ
 Query Match 100.0%; Score 29; DB 20; Length 6;
 Best Local Similarity 100.0%; Pred. No. 4.3e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMD 6
 DB 1 eellmd 6

RESULT 2
 AAY42738
 ID AAY42738 standard; peptide: 33 AA.

AC AAY42738;

DT 20-DEC-1999 (first entry)

DE Human alpha-1-antitrypsin C-terminal peptide 6.

XX Alpha-1-antitrypsin; fragment: cholesterol; cleavage:

KW low density lipoprotein; LDL; LDL receptor; hypercholesterolaemia;

KW atherosclerosis; gallstone.

OS Synthetic.

OS Homo sapiens.

FM Key Location/Qualifiers
 FT Modified-site 1 /note= "Optionally N-terminally acetylated, tosylated,
 FT myristoylated, benzoylated or carbobenzoxyated."

PN WO9945940-A1.

PD 16-SEP-1999.

PF 10-MAR-1999; 99WO-US05139.

PR 12-MAR-1998; 98US-0038935.

PA (UUYI-) UNIV VIRGINIA COMMONWEALTH.

PI Wright HT, Jancloauskiene S;

DR WPI, 1999-590819/50.

XX Lowering cholesterol levels in a patient using hypocholesterolemic
 PT peptide -

PS Disclosure; Page 5; 28pp; English.

XX This sequence represents human alpha-1-antitrypsin
 CC C-terminal peptide fragment 6. Alpha-1-antitrypsin is a
 CC known inhibitor of serine proteases, but loses its inhibitory activity
 CC due to a change in tertiary structure when cleaved by proteases such
 CC as leukocyte elastase. The cleaved alpha-1-antitrypsin molecules are
 CC cleared from the circulation through receptors in the liver and is
 CC accompanied by a depletion of extracellular cholesterol. The cause of
 CC this cholesterol depletion is due to an increase in the number of low
 CC density lipoprotein (LDL) receptors in liver cells which take up the

CC LDL cholesterol complex. This invention takes advantage of the fact
 CC that there is an increase in LDL receptor levels induced by the presence
 CC of cleaved alpha-1-antitrypsin and its derivatives, including C-terminal
 CC peptide fragments. The C-terminal peptide fragments (AAY42733-142749)
 CC can be used to reduce the levels of LDL cholesterol in a patient and can
 CC be used to treat a wide variety of disorders, including atherosclerosis,
 CC hypercholesterolaemia and gallstones. As the peptides are derived from a
 CC naturally occurring human serum protein, they should not produce immune
 CC side effects.

XX Sequence 33 AA;
 SQ
 Query Match 100.0%; Score 29; DB 20; Length 33;
 Best Local Similarity 100.0%; Pred. No. 7.7;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMD 6
 DB 4 eellmd 9

RESULT 3
 AAB92200
 ID AAB92200 standard; peptide: 33 AA.

AC AAB92200;

DT 22-JUN-2001 (first entry)

DE Thrombin inhibitor peptide SEQ ID NO:1376.

XX Protection: endogenous therapeutic peptide; peptidase: conjugation;

KW blood component; modification: succinimide; maleimido group; amino;

KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.

OS Synthetic.

OS Homo sapiens.

PN WO200069900-A2.

PD 23-NOV-2000.

PF 17-MAY-2000; 2000WO-US13576.

PR 17-MAY-1999; 99US-0134406.

PR 10-SEP-1999; 99US-0153406.

PR 15-OCT-1999; 99US-0159783.

PA (CONJ-) CONJUCHEM INC.

PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudau K;

DR WPI, 2001-112059/12.

XX Modifying and attaching therapeutic peptides to albumin prevents
 PT peptidase degradation, useful for increasing length of in vivo activity

PS Disclosure; Page 646; 733pp; English.

XX The present invention describes a modified therapeutic peptide (I)
 CC comprising a therapeutically active amino acid region (III) and a
 CC reactive group (II) (e.g., succinimide and maleimido groups) attached to
 CC a less therapeutically active amino acid region (IV), which covalently
 CC bonds with amino/hydroxyl/thiol groups on blood components to form a
 CC peptidase stabilized therapeutic peptide composed of 3-50 amino acids.
 CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
 CC factors and neurotransmitters, to protect them from peptidase activity
 CC in vivo for the treatment of various disorders. Endogenous therapeutic
 CC peptides are not suitable as drug candidates as they require frequent
 CC administration due to rapid degradation by peptidases in the body.
 CC Modifying and attaching therapeutic peptides to albumin prevents or

CC reduces the action of peptidases to increase length of activity (half
CC life) and specificity as bonding to large molecules decreases
CC intracellular uptake and interference with physiological processes.
CC AAB90829 to AAB92441 represent peptides which can be used in the
CC exemplification of the present invention.

XX Sequence 33 AA;

Query Match 100.0%; Score 29; DB 22; Length 33;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ERIIMD 6
|||||
Db 4 eelind 9

RESULT 4

AAG32041 standard; Protein: 220 AA.

XX AAG32041;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 38581.

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

XX Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126789.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134770.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140655.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142124.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144326.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.

PR	04-AUG-1999;	99US-0147204.	PR	29-OCT-1999;	99US-0162142.
PR	04-AUG-1999;	99US-0147302.	PR	29-OCT-1999;	99US-0162142.
PR	05-AUG-1999;	99US-0147192.	PR	29-OCT-1999;	99US-0162142.
PR	05-AUG-1999;	99US-0147260.	PR	29-OCT-1999;	99US-0162142.
PR	06-AUG-1999;	99US-0147303.	PR	29-OCT-1999;	99US-0162142.
PR	06-AUG-1999;	99US-0147416.	PR	29-OCT-1999;	99US-0162142.
PR	09-AUG-1999;	99US-0147493.	PR	29-OCT-1999;	99US-0162142.
PR	09-AUG-1999;	99US-0147935.	PR	29-OCT-1999;	99US-0162142.
PR	10-AUG-1999;	99US-0148171.	PR	29-OCT-1999;	99US-0162142.
PR	11-AUG-1999;	99US-0148341.	PR	29-OCT-1999;	99US-0162142.
PR	12-AUG-1999;	99US-0148341.	PR	29-OCT-1999;	99US-0162142.
PR	13-AUG-1999;	99US-0148565.	PR	29-OCT-1999;	99US-0162142.
PR	13-AUG-1999;	99US-0148565.	PR	29-OCT-1999;	99US-0162142.
PR	16-AUG-1999;	99US-0149368.	PR	29-OCT-1999;	99US-0162142.
PR	17-AUG-1999;	99US-0149175.	PR	29-OCT-1999;	99US-0162142.
PR	18-AUG-1999;	99US-0149426.	PR	29-OCT-1999;	99US-0162142.
PR	20-AUG-1999;	99US-0149723.	PR	29-OCT-1999;	99US-0162142.
PR	20-AUG-1999;	99US-0149723.	PR	29-OCT-1999;	99US-0162142.
PR	23-AUG-1999;	99US-0149902.	PR	29-OCT-1999;	99US-0162142.
PR	23-AUG-1999;	99US-0149902.	PR	29-OCT-1999;	99US-0162142.
PR	25-AUG-1999;	99US-0150560.	PR	29-OCT-1999;	99US-0162142.
PR	26-AUG-1999;	99US-0150884.	PR	29-OCT-1999;	99US-0162142.
PR	27-AUG-1999;	99US-0151065.	PR	29-OCT-1999;	99US-0162142.
PR	27-AUG-1999;	99US-0151065.	PR	29-OCT-1999;	99US-0162142.
PR	30-AUG-1999;	99US-0151303.	PR	29-OCT-1999;	99US-0162142.
PR	31-AUG-1999;	99US-0151303.	PR	29-OCT-1999;	99US-0162142.
PR	01-SEP-1999;	99US-0151930.	PR	29-OCT-1999;	99US-0162142.
PR	07-SEP-1999;	99US-0152363.	PR	29-OCT-1999;	99US-0162142.
PR	10-SEP-1999;	99US-0153070.	PR	29-OCT-1999;	99US-0162142.
PR	13-SEP-1999;	99US-0153758.	PR	29-OCT-1999;	99US-0162142.
PR	15-SEP-1999;	99US-0154018.	PR	29-OCT-1999;	99US-0162142.
PR	16-SEP-1999;	99US-0154039.	PR	29-OCT-1999;	99US-0162142.
PR	20-SEP-1999;	99US-0154779.	PR	29-OCT-1999;	99US-0162142.
PR	22-SEP-1999;	99US-0155138.	PR	29-OCT-1999;	99US-0162142.
PR	23-SEP-1999;	99US-0155488.	PR	29-OCT-1999;	99US-0162142.
PR	24-SEP-1999;	99US-0155653.	PR	29-OCT-1999;	99US-0162142.
PR	28-SEP-1999;	99US-0156458.	PR	29-OCT-1999;	99US-0162142.
PR	29-SEP-1999;	99US-0156458.	PR	29-OCT-1999;	99US-0162142.
PR	04-OCT-1999;	99US-0157117.	PR	29-OCT-1999;	99US-0162142.
PR	05-OCT-1999;	99US-0157533.	PR	29-OCT-1999;	99US-0162142.
PR	06-OCT-1999;	99US-0157865.	PR	29-OCT-1999;	99US-0162142.
PR	07-OCT-1999;	99US-0158023.	PR	29-OCT-1999;	99US-0162142.
PR	08-OCT-1999;	99US-0158232.	PR	29-OCT-1999;	99US-0162142.
PR	12-OCT-1999;	99US-0158369.	PR	29-OCT-1999;	99US-0162142.
PR	13-OCT-1999;	99US-0158923.	PR	29-OCT-1999;	99US-0162142.
PR	13-OCT-1999;	99US-0158923.	PR	29-OCT-1999;	99US-0162142.
PR	14-OCT-1999;	99US-0159332.	PR	29-OCT-1999;	99US-0162142.
PR	14-OCT-1999;	99US-0159332.	PR	29-OCT-1999;	99US-0162142.
PR	14-OCT-1999;	99US-0159330.	PR	29-OCT-1999;	99US-0162142.
PR	14-OCT-1999;	99US-0159331.	PR	29-OCT-1999;	99US-0162142.
PR	14-OCT-1999;	99US-0159637.	PR	29-OCT-1999;	99US-0162142.
PR	14-OCT-1999;	99US-0159637.	PR	29-OCT-1999;	99US-0162142.
PR	18-OCT-1999;	99US-0159584.	PR	29-OCT-1999;	99US-0162142.
PR	18-OCT-1999;	99US-0159584.	PR	29-OCT-1999;	99US-01

Db 42 eelind 47

RESULT 6

AAG32039

17-OCT-2000 (first entry)

Arabidopsis thaliana protein fragment SEQ ID NO: 38579.

Protein identification: signal transduction pathway; metabolic pathway;

hybridisation assay; genetic mapping; gene expression control; promoter;

termination sequence.

Arabidopsis thaliana.

EP103405-A2.

06-SEP-2000.

25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-012578.
XX 25-MAR-1999; 99US-0126284.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128274.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130821.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0132863.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 18-MAY-1999; 99US-0134370.
XX 19-MAY-1999; 99US-0134768.
XX 20-MAY-1999; 99US-0134941.
XX 21-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136382.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 14-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142877.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145146.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.

PR 17-AUG-1999; 9905-0149175.
 PR 18-AUG-1999; 9905-0149426.
 PR 20-AUG-1999; 9905-0149722.
 PR 20-AUG-1999; 9905-0149723.
 PR 20-AUG-1999; 9905-0149929.
 PR 23-AUG-1999; 9905-0149902.
 PR 23-AUG-1999; 9905-0149930.
 PR 25-AUG-1999; 9905-0150366.
 PR 26-AUG-1999; 9905-0150884.
 PR 27-AUG-1999; 9905-0151065.
 PR 27-AUG-1999; 9905-0151066.
 PR 27-AUG-1999; 9905-0151080.
 PR 30-AUG-1999; 9905-0151303.
 PR 31-AUG-1999; 9905-0151338.
 PR 01-SEP-1999; 9905-0151930.
 PR 07-SEP-1999; 9905-0152363.
 PR 10-SEP-1999; 9905-0153070.
 PR 13-SEP-1999; 9905-0153756.
 PR 13-SEP-1999; 9905-0154018.
 PR 16-SEP-1999; 9905-0154039.
 PR 20-SEP-1999; 9905-0154779.
 PR 22-SEP-1999; 9905-0155139.
 PR 23-SEP-1999; 9905-0155486.
 PR 24-SEP-1999; 9905-0155659.
 PR 28-SEP-1999; 9905-0156458.
 PR 29-SEP-1999; 9905-0156596.
 PR 04-OCT-1999; 9905-0157117.
 PR 05-OCT-1999; 9905-0157753.
 PR 06-OCT-1999; 9905-0157865.
 PR 07-OCT-1999; 9905-0158029.
 PR 08-OCT-1999; 9905-0158232.
 PR 12-OCT-1999; 9905-0158369.
 PR 13-OCT-1999; 9905-0159293.
 PR 13-OCT-1999; 9905-0159294.
 PR 13-OCT-1999; 9905-0159295.
 PR 14-OCT-1999; 9905-0159329.
 PR 14-OCT-1999; 9905-0159330.
 PR 14-OCT-1999; 9905-0159331.
 PR 14-OCT-1999; 9905-0159637.
 PR 14-OCT-1999; 9905-0159638.
 PR 18-OCT-1999; 9905-0159584.
 PR 21-OCT-1999; 9905-0160741.
 PR 21-OCT-1999; 9905-0160767.
 PR 21-OCT-1999; 9905-0160768.
 PR 21-OCT-1999; 9905-0160770.
 PR 21-OCT-1999; 9905-0160814.
 PR 21-OCT-1999; 9905-0160815.
 PR 22-OCT-1999; 9905-0160980.
 PR 22-OCT-1999; 9905-0160981.
 PR 22-OCT-1999; 9905-0160989.
 PR 23-OCT-1999; 9905-0161404.
 PR 23-OCT-1999; 9905-0161405.
 PR 25-OCT-1999; 9905-0161406.
 PR 26-OCT-1999; 9905-0161359.
 PR 26-OCT-1999; 9905-0161360.
 PR 26-OCT-1999; 9905-0161361.
 PR 28-OCT-1999; 9905-0161920.
 PR 28-OCT-1999; 9905-0161922.
 PR 28-OCT-1999; 9905-0161993.
 PR 29-OCT-1999; 9905-0162142.

Query Match 100.0%; Score 29; DB 21; Length 284;
 Best Local Similarity 100.0%; Pred. No. 64;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIND 6
 Db 90 eelind 95

RESULT 7
 AAR11756
 ID AAR11756 standard; Protein; 379 AA.

XX AAR11756;
 AC 10-JUL-1991 (first entry)
 DT Human PAI-1 P1-P1' ATIII mutant.
 DE plasminogen activator inhibitor; antithrombin III; vitronectin;
 KW fibrinolysis; thrombolysis; serpin.
 OS Homo sapiens.
 PN MO9105048-A.
 PD 18-APR-1991.
 XX 03-OCT-1990; 90MO-NL00145.
 PF 03-OCT-1989; 89NL-0002454.
 PR (STCL) STICHT INST AND BLOEDTRAN.
 PA Pannekoek H;
 PI WPI; 1991-132855/18.
 DR New plasminogen activator inhibitor 1 (PAI-1) mutants - comprise
 XX PT region of antithrombin III (ATIII), useful in fibrinolytic-
 PT thrombolytic therapy
 XX
 XX Claim 2; Page 24; 32pp; English.
 CC This sequence comprises mature PAI-1 in which Met(347) from the
 CC reactive centre (SGTVASSTFAVTSARAPETIMD) has been replaced by
 CC Ser from the corresponding site of ATIII. The mutant PAI-1 is used
 CC to prevent reocclusion. It binds to vitronectin to act as a strong
 CC thrombin inhibitor, but is much less effective as a PA inhibitor.
 CC See also AAR11755 and AAR12162-3.
 SO Sequence 379 AA;

Query Match 100.0%; Score 29; DB 12; Length 379;
 Best Local Similarity 100.0%; Pred. No. 84;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIND 6
 Db 350 eelind 355

RESULT 8
 AAR12162
 ID AAR12162 standard; Protein; 379 AA.
 XX AAR12162;
 AC 10-JUL-1991 (first entry)
 DT Human PAI-1 P2-P2' ATIII mutant.
 DE plasminogen activator inhibitor; antithrombin III; vitronectin;
 KW fibrinolysis; thrombolysis; serpin.
 OS Homo sapiens.
 PN MO9105048-A.
 PD 18-APR-1991.
 XX 03-OCT-1990; 90MO-NL00145.
 PF 03-OCT-1989; 89NL-0002454.

XX (STCL) STICHT INST AND BLOEDTRAN.
 PA Pannekoek H;
 PI WPI; 1991-132855/18.
 XX
 DR
 XX
 PT New plasminogen activator inhibitor 1 (PAI-1) mutants - comprise
 PT region of antithrombin III (ATIII), useful in fibrinolytic-
 PT thrombolytic therapy
 PS
 XX Claim 3; Page 24; 32pp: English.
 XX
 CC This sequence comprises mature PAI-1 in which residues 345-348
 CC (i.e. AlaArgMetAla) in the reactive centre has been replaced by
 CC GlyArgSerLeu from the corresponding site of ATIII. The mutant PAI-1
 CC is used to prevent reocclusion. It binds to vitronectin to act as a
 CC strong thrombin inhibitor, but is much less effective as a PA
 CC inhibitor.
 CC See also AAR11755-6 and AAR12163.
 CC
 SO Sequence 379 AA;
 SO

Query Match 100.0%; Score 29; DB 12; Length 379;
 Best Local Similarity 100.0%; Pred. No. 84;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EEIIMD 6
 |||||
 DB 350 eelind 355

RESULT 9
 AAR12163
 ID AAR12163 standard; Protein; 379 AA.
 XX
 AC AAR12163;
 XX
 DT 10-JUL-1991 (first entry)
 XX
 DE Human PAI-1 P3-P3' ATIII mutant.
 XX
 KW plasminogen activator inhibitor; antithrombin III; vitronectin;
 KW fibrinolysis; thrombolysis; serpin.
 XX
 OS Homo sapiens.
 XX
 PN WO9105048-A.
 XX
 PD 18-APR-1991.
 XX
 PF 03-OCT-1990; 90WO-NL00145.
 XX
 PR 03-OCT-1989; 89NL-0002454.
 XX
 PA (STCL) STICHT INST AND BLOEDTRAN.
 XX
 PI Pannekoek H;
 XX
 DR WPI; 1991-132855/18.
 XX
 PT New plasminogen activator inhibitor 1 (PAI-1) mutants - comprise
 PT region of antithrombin III (ATIII), useful in fibrinolytic-
 PT thrombolytic therapy
 PS
 XX Claim 4; Page 24; 32pp: English.
 XX
 CC This sequence comprises mature PAI-1 in which residues 344-349
 CC (i.e. SerAlaArgMetAlaPro) in the reactive centre has been replaced
 CC by AlaGlyArgSerLeu from the corresponding site of ATIII. The
 CC mutant PAI-1 is used to prevent reocclusion. It binds to vitronectin
 CC to act as a strong thrombin inhibitor, but is much less effective

CC as a PA inhibitor.
 CC See also AAR11755-6 and AAR12162.
 XX
 SO Sequence 379 AA;
 SO

Query Match 100.0%; Score 29; DB 12; Length 379;
 Best Local Similarity 100.0%; Pred. No. 84;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EEIIMD 6
 |||||
 DB 350 eelind 355

RESULT 10
 AAM26710
 ID AAM26710 standard; Protein; 379 AA.
 XX
 AC AAM26710;
 XX
 DT 14-APR-1998 (first entry)
 XX
 DE Plasminogen activator inhibitor mutant P1Aa (R346A).
 XX
 DE Plasminogen activator inhibitor type 1; PAI-1; human;
 KW elastase; inhibitor; vitronectin; cell attachment; cell migration;
 KW cell proliferation; emphysema; adult respiratory distress syndrome;
 KW acute lung inflammation; alpha 1-antitrypsin deficiency;
 KW cystic fibrosis; atopic dermatitis; pancreatitis;
 KW periodontal disease; arthritis; HIV; atherosclerosis; restenosis;
 KW neointima; fibrosis; wound healing; tumour; metastasis; psoriasis;
 KW thrombosis; angiogenesis; therapy.
 XX
 OS Homo sapiens.
 OS
 OS Synthetic.
 XX
 PN WO9739028-A1.
 XX
 PD 23-OCT-1997.
 XX
 PF 11-APR-1997; 97WO-US06071.
 XX
 PR 12-APR-1996; 96US-0015299.
 XX
 PA (AMNA-) AMERICAN NAT RED CROSS.
 XX
 PA Lawrence DA, Stefansson SP;
 XX
 DR WPI; 1997-526399/48.
 XX
 PT Plasminogen activator-inhibitor type I mutant inhibits elastase - or
 PT has high affinity for vitronectin for therapeutic inhibition of
 PT elastase or vitronectin-mediated cell attachment, migration etc.
 XX
 PS Claim 4; Page -: 144pp: English.
 XX
 CC This polypeptide comprises a preferred mutant, P1Aa (R346A), of
 CC human plasminogen activator inhibitor (PAI-1) mature protein (see
 CC also AAM31587). It has an Ala for Arg substitution at amino acid
 CC residue 346, i.e. in the reactive centre loop of the protein. The
 CC mutant can be obtained by site-directed mutagenesis of PAI-1 DNA
 CC (see AAG7303) and expression in a host cell. Claimed PAI-1 mutants
 CC (see AAM26710-25) have at least one amino acid substitution at
 CC position 343-350, especially at position 343 and/or 346, and may
 CC have additional stabilising mutations comprising between one and
 CC four of His-150, Thr-154, Leu-319 and Ile-354, and one or more of
 CC Arg-333, Arg-335, Gly-331, Ile-372 and Leu-91. The mutants are
 CC characterised by being resistant to inactivation by elastase,
 CC plasminogen activator, plasmin, cathepsin G, chymase, gelatinases A
 CC and B, stromelysin and collagenase, and by having a high affinity
 CC for vitronectin (Vn) and are used, in vivo or in vitro, to inhibit
 CC pathological elastase activity (specifically in emphysema, adult

XX

XX 5

CC This polypeptide comprises a preferred mutant, P1Gly (R346S), of
CC human plasminogen activator inhibitor (PAI-1) mature protein (see
CC also A631587). It has a Gly for Arg substitution at amino acid
CC residue 346, i.e. in the reactive centre loop of the protein. The
CC mutant can be obtained by site-directed mutagenesis of PAI-1 DNA
CC (see AAT97301) and expression in a host cell. Claimed PAI-1 mutants
CC (see AAW26710-25) have at least one amino acid substitution at
CC position 343-350, especially at position 343 and/or 346, and may
CC have additional stabilising mutations comprising between one and
CC four of His-150, Thr-154, Leu-319 and Ile-354, and one or more of
CC Arg-333, Arg-335, Gly-331, Ile-372 and Leu-31. The mutants are
CC characterised by being resistant to inactivation by elastase,
CC plasminogen activator, plasmin, cathepsin G, chymase, gelatinases A
CC and B, stromelysin and collagenase, and by having a high affinity
CC for vitronectin (Vn) and are used, in vivo or in vitro, to inhibit
CC pathological elastase activity (specifically in emphysema, adult
CC respiratory distress syndrome, acute lung inflammation, congenital
CC alpha 1 antitrypsin deficiency, cystic fibrosis, atopic dermatitis,
CC pancreatitis, periodontal disease, arthritis and HIV infection).
CC They also inhibit Vn-stimulated cell attachment, migration and/or
CC migration-induced proliferation (specifically in atherosclerosis,
CC post-angioplasty restenosis, neointima formation after vascular
CC trauma, vascular graft restenosis, fibrosis (associated with
CC chronic inflammation or chemotherapy, or of the lung), wound
CC healing with scarring and fibrosis), growth/invasion of primary
CC tumours or metastases, psoriasis, deep vein thrombosis and any
CC condition involving pathological angiogenesis (all claimed) and
CC the mutants are resistant to cleavage by the proteases to which
CC they bind, they are more effective therapeutic agents. They also
CC promote uptake and clearance of elastase (or its complex with PAI-1)
CC by receptors.
CC (NB, the amino acid sequence of the mutant protein was obtained by
CC adaptation of the wild-type sequence provided in Fig 4A of the
CC specification).

QX Sequence 379 AA:

QY	1	EEIIMD	6
Db	350	eelImd	355
RESULT 13			
ID	AAW26713		
	AAW26713	standard; Protein: 379 AA.	
XX	AC	AAW26713:	
XX	AC		
DT	14-Apr-1998	(first entry)	
XX			
DE	Plasminogen activator inhibitor mutant PIAsp (R346D).		
XX			

PF 11-APR-1997; 97WO-US06071.
 XX
 PR 12-APR-1996; 96US-0015299.
 XX
 PR (AMNA-) AMERICAN NAT RED CROSS.
 PA
 XX
 PI Lawrence DA, Stefansson SP;
 XX
 DR WPI: 1997-526399/48.
 XX
 PT plasminogen activator-inhibitor type I mutant inhibits elastase - or
 PT has high affinity for vitronectin, for therapeutic inhibition of
 PT elastase or vitronectin-mediated cell attachment, migration etc.
 XX
 XS Claim 4; Page -: 144pp; English.

CC This polypeptide comprises a preferred mutant, Plasp (R346p), of
CC human plasminogen activator inhibitor (PAI-1) mature protein (see
CC also AA031587). It has an Asp for Arg substitution at amino acid
CC residue 346, i.e. in the reactive centre loop of the protein. The
CC mutant can be obtained by site-directed mutagenesis of PAI-1 DNA
CC (see AA079303) and expression in a host cell. Claimed PAI-1 mutants
CC (see AA062610-25) have at least one amino acid substitution at:
CC position 343-350, especially at position 343 and/or 346, and may
CC have additional stabilising mutations comprising between one and
CC four of His-150, Asp-154, Leu-319 and Ile-351, and one or more of
CC Arg-333, Arg-335, Gly-331, Ile-372 and Leu-51. The mutants are
CC characterised by being resistant to inactivation by elastase,
CC plasminogen activator, plasmin, cathepsin G, chymase, gelatinases A
CC and B, stromelysin and collagenase, and by having a high affinity
CC for vitronectin (Vn) and are used, in vivo or in vitro, to inhibit
CC pathological elastase activity (specifically in lymphoma, adult
CC respiratory distress syndrome, acute lung inflammation, congenital
CC alpha-1-antitrypsin deficiency, cystic fibrosis, atopic dermatitis,
CC pancreatitis, periodontal disease, arthritis and HIV infection).
CC They also inhibit Vn-stimulated cell attachment, migration and/or
CC migration-induced proliferation (specifically in atherosclerosis,
CC post-angioplasty restenosis, neointima formation after vascular
CC trauma, vascular graft restenosis, fibrosis (associated with
CC chronic inflammation or chemotherapy, or of the lung), wound
CC healing with scarring and fibrosis, growth/invasion of primary
CC tumours or metastases, psoriasis, deep vein thrombosis and any
CC condition involving pathological angiogenesis (all claimed). Since
CC the mutants are resistant to cleavage by the proteases to which
CC they bind, they are more effective therapeutic agents. They also
CC promote uptake and clearance of elastase (or its complex with PAI-1)
CC by receptors.
CC (NB. the amino acid sequence of the mutant protein was obtained by
CC adaptation of the wild-type sequence provided in Fig 4A of the
CC specification).
XX Sequence 379 AA;

Query Match	100.0%;	Score 29;	DB 18;	Length 379;
Best Local Similarity	100.0%;	Pred. No. 84;		
Matches	6;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0

QY 1 EEIIMD 6
|||||
DB 350 eellmd 355

RESULT 14
AAW26714
ID AAW26714 standard; Protein: 379 AA.
XX
AC AAW26714;
XX
DT 14-APR-1998 (first entry)
XX
DE Plasminogen activator inhibitor mutant P4Ala (V343A).
XX
KM Plasminogen activator inhibitor type 1; PAI-1; human;
KM elastase; inhibitor; vitronectin; cell attachment; cell migration;
KM cell proliferation; emphysema; adult respiratory distress syndrome;
KM acute lung inflammation; alpha 1-antitrypsin deficiency;
KM cystic fibrosis; atopic dermatitis; pancreatitis;
KM periodontal disease; arthritis; HIV; atherosclerosis; restenosis;
KM neointima; fibrosis; wound healing; tumour; metastasis; psoriasis;
KM thrombosis; angiogenesis; therapy.
XX
OS Homo sapiens.
OS Synthetic.
XX
PM W09739028-A1.
XX
PD 23-OCT-1997.
XX
PF 11-APR-1997; 97WO-US06071.
XX
PR 12-APR-1996; 96US-0015299.
XX
PA (AMNA-) AMERICAN NAT RED CROSS.
XX
PI Lawrence DA, Stefansson SP;
XX
DR WPI; 1997-526399/48.
XX
PT Plasminogen activator-inhibitor type I mutant inhibits elastase - or
PT has high affinity for vitronectin, for therapeutic inhibition of
PT elastase or vitronectin-mediated cell attachment, migration etc.
XX
PS Claim 4; Page -: 144pp; English.

This polypeptide comprises a preferred mutant, P4Ala (V343A), of human plasminogen activator inhibitor (PAI-1) mature protein (see also AAW31587). It has an Ala for Val substitution at amino acid residue 343, i.e. in the reactive centre loop of the protein. The mutant can be obtained by site-directed mutagenesis of PAI-1 DNA (see AAW97303) and expression in a host cell. Claimed PAI-1 mutants (see AAW26710-25) have at least one amino acid substitution at position 343-350, especially at position 343 and/or 346, and may have additional stabilising mutations comprising between one and four of His-150, Thr-154, Leu-319 and Ile-354, and one or more of Arg-333, Arg-335, Gly-331, Ile-372 and Leu-91. The mutants are characterised by being resistant to inactivation by elastase. CC plasminogen activator, plasmin, cathepsin G, chymase, gelatinases A and B, streptolysin and collagenase, and by having a high affinity for vitronectin (Vn) and are used, in vivo or in vitro, to inhibit pathological elastase activity (specifically in emphysema, adult respiratory distress syndrome, acute lung inflammation, congenital alpha 1-antitrypsin deficiency, cystic fibrosis, atopic dermatitis, pancreatitis, periodontal disease, arthritis and HIV infection). CC They also inhibit Vn-stimulated cell attachment, migration and/or CC migration-induced proliferation (specifically in atherosclerosis, CC post-angioplasty restenosis, neointima formation after vascular CC trauma, vascular graft restenosis, fibrosis (associated with CC chronic inflammation or chemotherapy, or of the lung), wound CC healing with scarring and fibrosis, growth/invasion of primary

CC tumours or metastases, psoriasis, deep vein thrombosis and any
CC condition involving pathological angiogenesis (all claimed). Since
CC the mutants are resistant to cleavage by the proteases to which
CC they bind, they are more effective therapeutic agents. They also
CC promote uptake and clearance of elastase (or its complex with PAI-1)
CC by receptors.
CC (NB, the amino acid sequence of the mutant protein was obtained by
CC adaptation of the wild-type sequence provided in Fig 4A of the
CC specification).
XX
SQ Sequence 379 AA;

Query Match 100.0%; Score 29; DB 18; Length 379;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEIIMD 6
|||||
DB 350 eellmd 355

RESULT 15
AAW26715
ID AAW26715 standard; Protein: 379 AA.
XX
AC AAW26715;
XX
DT 14-APR-1998 (first entry)
XX
DE Plasminogen activator inhibitor mutant P4Asp (V343D).
XX
KM Plasminogen activator inhibitor type 1; PAI-1; human;
KM elastase; inhibitor; vitronectin; cell attachment; cell migration;
KM cell proliferation; emphysema; adult respiratory distress syndrome;
KM acute lung inflammation; alpha 1-antitrypsin deficiency;
KM cystic fibrosis; atopic dermatitis; pancreatitis;
KM periodontal disease; arthritis; HIV; atherosclerosis; restenosis;
KM neointima; fibrosis; wound healing; tumour; metastasis; psoriasis;
KM thrombosis; angiogenesis; therapy.
XX
OS Homo sapiens.
OS Synthetic.
XX
PM W09739028-A1.
XX
PD 23-OCT-1997.
XX
PF 11-APR-1997; 97WO-US06071.
XX
PR 12-APR-1996; 96US-0015299.
XX
PA (AMNA-) AMERICAN NAT RED CROSS.
XX
PI Lawrence DA, Stefansson SP;
XX
DR WPI; 1997-526399/48.
XX
PT Plasminogen activator-inhibitor type I mutant inhibits elastase - or
PT has high affinity for vitronectin, for therapeutic inhibition of
PT elastase or vitronectin-mediated cell attachment, migration etc.
XX
PS Claim 4; Page -: 144pp; English.

This polypeptide comprises a preferred mutant, P4Asp (V343D), of human plasminogen activator inhibitor (PAI-1) mature protein (see also AAW31587). It has an Asp for Val substitution at amino acid residue 343, i.e. in the reactive centre loop of the protein. The mutant can be obtained by site-directed mutagenesis of PAI-1 DNA (see AAW97303) and expression in a host cell. Claimed PAI-1 mutants (see AAW26710-25) have at least one amino acid substitution at position 343-350, especially at position 343 and/or 346, and may have additional stabilising mutations comprising between one and

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:49:40 ; Search time 37.77 Seconds
(without alignments)
3.575 Million cell updates/sec

Title: US-09-544-665-3
Perfect score: 29
Sequence: 1 EELMD 6

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, NA:
1: /cgn2_6/pdata/2/1aa/5a_COMB.pep:*
2: /cgn2_6/pdata/2/1aa/5b_COMB.pep:*
3: /cgn2_6/pdata/2/1aa/6a_COMB.pep:*
4: /cgn2_6/pdata/2/1aa/6b_COMB.pep:*
5: /cgn2_6/pdata/2/1aa/PCUTUS_COMB.pep:*
6: /cgn2_6/pdata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	29	100.0	33 1	US-08-781-020-6 Sequence 6, Appl
2	29	100.0	33 4	US-09-038-935-6 Sequence 6, Appl
3	29	100.0	379 3	US-08-840-204-3 Sequence 3, Appl
4	29	100.0	390 1	US-08-121-714-6 Sequence 6, Appl
5	29	100.0	380 1	US-08-477-108A-6 Sequence 6, Appl
6	29	100.0	380 2	US-08-477-112-6 Sequence 6, Appl
7	29	100.0	390 5	PCY-US93-08322-6 Sequence 6, Appl
8	29	100.0	402 1	US-08-315-461-7 Sequence 7, Appl
9	29	100.0	402 3	US-08-948-997-4 Sequence 4, Appl
10	29	100.0	402 3	US-08-840-204-2 Sequence 2, Appl
11	29	100.0	402 4	US-09-348-817A-4 Sequence 4, Appl
12	29	100.0	402 4	US-09-026-408-3 Sequence 3, Appl
13	26	89.7	551 4	US-08-796-899-29 Sequence 158, App
14	25	86.2	471 4	US-08-961-083-158 Sequence 3, Appl
15	24	82.8	200 3	US-09-019-942-3 Sequence 3, Appl
16	24	82.8	226 3	US-09-176-657-3 Sequence 4, Appl
17	24	82.8	253 2	US-08-642-045B-4 Sequence 4, Appl
18	24	82.8	253 4	US-08-852-268-2 Sequence 2, Appl
19	24	82.8	362 4	US-09-196-520-4 Sequence 8, Appl
20	24	82.8	362 4	US-09-196-520-8 Sequence 2, Appl
21	24	82.8	365 4	US-09-196-520-2 Sequence 5, Appl
22	24	82.8	429 2	US-08-677-049-5 Sequence 3, Appl
23	24	82.8	586 3	US-08-964-268-3 Sequence 3, Appl
24	24	82.8	656 4	US-09-376-781-3 Sequence 4, Appl
25	24	82.8	657 3	US-09-111-085-4 Sequence 4, Appl
26	24	82.8	657 4	US-09-376-781-4 Sequence 10, Appl
27	24	82.8	803 1	US-08-158-232-10 Sequence 10, Appl

28	24	82.8	803 1	US-08-304-626-10 Sequence 10, Appl
29	24	82.8	803 1	US-08-316-301A-12 Sequence 12, Appl
30	24	82.8	803 2	US-08-611-928-10 Sequence 10, Appl
31	24	82.8	803 3	US-09-173-891-10 Sequence 10, Appl
32	24	82.8	803 4	US-09-076-137-12 Sequence 12, Appl
33	24	82.8	803 5	PCT-US92-03624-12 Sequence 12, Appl
34	24	82.8	855 2	US-08-482-090-12 Sequence 8, Appl
35	24	82.8	856 2	US-08-481-700B-8 Sequence 16, Appl
36	24	82.8	856 2	US-09-007-393-16 Sequence 4, Appl
37	24	82.8	863 1	US-08-325-347-4 Sequence 2, Appl
38	24	82.8	1176 1	US-08-434-823-2 Sequence 3, Appl
39	24	82.8	1176 1	US-08-457-366-2 Sequence 16, Appl
40	24	82.8	1724 1	US-08-325-547-3 Sequence 8, Appl
41	23	79.3	8 2	US-08-467-046-16 Sequence 20, Appl
42	23	79.3	14 2	US-08-426-627-20 Sequence 8, Appl
43	23	79.3	154 1	US-08-033-857A-8 Sequence 8, Appl
44	23	79.3	164 1	US-08-374-983A-8 Sequence 8, Appl
45	23	79.3	164 1	US-08-374-983A-8 Sequence 8, Appl

ALIGNMENTS

```

RESULT 1
US-08-781-020-6
; Sequence 6, Application US/08781020
; Patent No. 5792749
GENERAL INFORMATION:
APPLICANT: Wright, H. Tonie
TITLE OF INVENTION: METHOD AND COMPOSITION FOR LOWERING LOW
DENSITY LIPOPROTEIN CHOLESTEROL
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESSES:
ADDRESS: WHITHAM, CURTIS, WHITHAM & MCGINN
STREET: Reston Intl. Center, 11800 Sunrise Valley Dr.,
STREET: Suite 900
CITY: Reston
STATE: VA
COUNTRY: USA
ZIP: 20191
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,020
FILING DATE: 09-JAN-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Whitham, Michael E.
REGISTRATION NUMBER: 32,635
REFERENCE/DOCKET NUMBER: VCU-16
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-391-9035
TELEFAX: 703-391-9035
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: Peptide
US-08-781-020-6

Query Match 100.0%; Score 29; DB 1; Length 33;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 EELMD 6
|||||
DB 4 EELMD 9

```


RESULT 2
US-09-038-935-6
; Sequence 6, Application US/09038935
; Patent No. 6150332
; GENERAL INFORMATION:
; APPLICANT: Wright, H. Tonle
; APPLICANT: Janciuskiene, Sabina
; TITLE OF INVENTION: METHOD AND COMPOSITION FOR LOWERING LOW DENSITY
; TITLE OF INVENTION: LIPOPROTEIN CHOLESTEROL
; FILE REFERENCE: 294022ba
; CURRENT APPLICATION NUMBER: US/09/038,935
; CURRENT FILING DATE: 1998-03-12
; EARLIER APPLICATION NUMBER: 08/781,020
; EARLIER FILING DATE: 1997-01-19
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-038-935-6

Query Match 100.0%; Score 29; DB 4; Length 33;
Best Local Similarity 100.0%; Pred. No. 3,1;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EELIMD 6
DB 4 EELIMD 9

RESULT 3
US-08-840-204-3
; Sequence 3, Application US/08840204
; Patent No. 6103498
; GENERAL INFORMATION:
; APPLICANT: LAWRENCE, DANIEL A.
; APPLICANT: STEFANSSON, STEINGRIMUR P.
; TITLE OF INVENTION: MUTANT PLASMINOGEN ACTIVATOR-INHIBITOR
; TITLE OF INVENTION: TYPE 1 (PAI-1) AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FORSEER
; STREET: 2000 PENNSYLVANIA AVENUE, NW
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/840,204
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: LIVNAT, SHMUEL
; REGISTRATION NUMBER: 33,949
; REFERENCE/DOCKET NUMBER: 30807-20004.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 822-0168
; TELEX: 90-4030 MRSNFORSMH
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

TOPOLOGY: linear
US-08-840-204-3

Query Match 100.0%; Score 29; DB 3; Length 379;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EELIMD 6
DB 350 EELIMD 355

RESULT 4
US-08-121-714-6
; Sequence 6, Application US/08121714
; Patent No. 5470970
; GENERAL INFORMATION:
; APPLICANT: Sager, Ruth
; TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH
; TITLE OF INVENTION: TUMOR SUPPRESSING ACTIVITY
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 555X
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,714
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/938,823
; FILING DATE: 09/01/92
; APPLICATION NUMBER: 07/844,296
; FILING DATE: 02/28/92
; APPLICATION NUMBER: 07/662,216
; FILING DATE: 02/28/91
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/072001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 390
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-121-714-6

Query Match 100.0%; Score 29; DB 1; Length 390;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EELIMD 6
DB 373 EELIMD 378

RESULT 5
US-08-477-108A-6

Sequence 6, Application US/08477108A
Patent No. 5801001
GENERAL INFORMATION:
APPLICANT: Sager, Ruth
APPLICANT: Zou, Zhigang
APPLICANT: Antosiewicz, Anthony
TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: Wordperfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,108A
FILING DATE: June 7, 1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/121,714
FILING DATE: 09/01/93
APPLICATION NUMBER: 07/938,823
FILING DATE: 09/01/92
APPLICATION NUMBER: 07/844,296
FILING DATE: 02/28/92
APPLICATION NUMBER: 07/662,216
FILING DATE: 02/28/91
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 06570/002002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 390
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-477-108A-6

Query Match 100.0%; Score 29; DB 1; Length 390;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EIIIMD 6
DB 373 EIIIMD 378
RESULT 6
US-08-477-112-6
Sequence 6, Application US/08477112
Patent No. 5905023
GENERAL INFORMATION:
APPLICANT: Sager, Ruth
APPLICANT: Zou, Zhigang
TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston

STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: Wordperfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,112
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/121,714
FILING DATE: 09/01/93
APPLICATION NUMBER: 07/938,823
FILING DATE: 09/01/92
APPLICATION NUMBER: 07/844,296
FILING DATE: 02/28/92
APPLICATION NUMBER: 07/662,216
FILING DATE: 02/28/91
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 06570/002003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 390
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-477-112-6

Query Match 100.0%; Score 29; DB 2; Length 390;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EIIIMD 6
DB 373 EIIIMD 378
RESULT 7
PCT-US93-08322-6
Sequence 6, Application PCT/US9308322
GENERAL INFORMATION:
APPLICANT: Sager, Ruth
TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH TUMOR SUPPRESSING ACTIVITY
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: Wordperfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/08322
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/938,823
FILING DATE: 09/01/92

APPLICATION NUMBER: 07/844,296
FILING DATE: 02/28/92
APPLICATION NUMBER: 07/662,216
FILING DATE: 02/28/91
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 00530/072001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 390
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
PCT-US93-08322-6

Query Match: 100.0%; Score 29; DB 5; Length 390;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EITMD 6
Db 373 EITMD 378

RESULT 8
US-08-315-461-7
Sequence 7, Application US/08315461
Patent No. 5639726
GENERAL INFORMATION:
APPLICANT: Lawrence, Daniel A.
APPLICANT: Ginsburg, David
APPLICANT: Shore, Joseph D.
APPLICANT: Fay, William P.
APPLICANT: Olson, Steven T.
APPLICANT: Francis-Chmura, Anne-Marie
APPLICANT: Daniel T. Eltzman
TITLE OF INVENTION: Peptide Mediated Enhancement of
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Arnold, White & Durkee
STREET: P. O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/315,461
FILING DATE: Concurrently Herewith
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UMIC:006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 402 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-315-461-7

Query Match: 100.0%; Score 29; DB 1; Length 402;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EITMD 6
Db 373 EITMD 378

RESULT 9
US-08-948-997-4
Sequence 4, Application US/08948997
Patent No. 6008020
GENERAL INFORMATION:
APPLICANT: HASTINGS, GREGG
APPLICANT: COLEMAN, TIM
APPLICANT: LAWRENCE, DANIEL
TITLE OF INVENTION: BRAIN-ASSOCIATED INHIBITOR OF
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESSES:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/948,997
FILING DATE: Oct-10-97
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: A. ANDERS BROOKES
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF336
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 402 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-948-997-4

Query Match: 100.0%; Score 29; DB 3; Length 402;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EITMD 6
Db 373 EITMD 378

RESULT 10
US-08-840-204-2
Sequence 2, Application US/08840204
Patent No. 6103498
GENERAL INFORMATION:

APPLICANT: LAWRENCE, DANIEL A.
TITLE OF INVENTION: MUTANT PLASMINOGEN ACTIVATOR-INHIBITOR
TITLE OF INVENTION: TYPE 1 (PAT-1) AND USES THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSER: MORRISON & FOERSTER
STREET: 2000 PENNSYLVANIA AVENUE, NW
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20006-1812

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/840.204
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: LIVNAT, SHMUEL
REGISTRATION NUMBER: 33,949
REFERENCE/DOCKET NUMBER: 30807-20004.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 822-0168
TELEX: 90-4030 MRSNFOERMSH
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 402 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-840-204-2

Query Match
Best Local Similarity 100.0%; Score 29; DB 3; Length 402;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EITMD 6
DB 373 EITMD 378

RESULT 11
US-09-348-817A-4
Sequence 4, Application US/09348817A
Patent No. 6191260
GENERAL INFORMATION:
APPLICANT: Hastings et al.
TITLE OF INVENTION: Brain-Associated Inhibitor of Tissue-Type Plasminogen
FILE REFERENCE: PF336D1
CURRENT APPLICATION NUMBER: US/09/348,817A
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: 08/948,997
PRIOR FILING DATE: 1997-10-10
PRIOR APPLICATION NUMBER: 60/028,117
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 4
LENGTH: 402
TYPE: PRT
ORGANISM: Bos taurus
US-09-348-817A-4

Query Match
Best Local Similarity 100.0%; Score 29; DB 4; Length 402;
Matches 100.0%; Pred. No. 38;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EITMD 6
DB 373 EITMD 378

RESULT 12
US-09-026-408-3
Sequence 3, Application US/09026408
Patent No. 630338
GENERAL INFORMATION:
APPLICANT: NI et al.
TITLE OF INVENTION: PANCREAS-DERIVED PLASMINOGEN ACTIVATOR
TITLE OF INVENTION: INHIBITOR
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVENUE, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/026,408
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/934,011
FILING DATE: 15-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/024,056
FILING DATE: 16-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.0300002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 402 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-026-408-3

Query Match
Best Local Similarity 100.0%; Score 29; DB 4; Length 402;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EITMD 6
DB 373 EITMD 378

RESULT 13
US-08-796-899-29
Sequence 29, Application US/08796899
Patent No. 6160202
GENERAL INFORMATION:
APPLICANT: BUSTOS, Mauricio M
TITLE OF INVENTION: MODIFICATION OF SEED CROPS WITH
TITLE OF INVENTION: TRANSCRIPTION FACTORS

NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 2213-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/796,899
FILING DATE: 06-FEB-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/319,544
FILING DATE: 07-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Meuth, Donna M
REGISTRATION NUMBER: 36,607
REFERENCE/DOCKET NUMBER: 028754-005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 551 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-796-899-29

Query Match
Best Local Similarity 89.7%; Score 26; DB 4; Length 551;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EIIIMD 6
Db 380 EIIIMD 385

RESULT 14
US-08-961-083-158
Sequence 158, Application US/08961083
Patent No. 6159469
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,083
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 158:
SEQUENCE CHARACTERISTICS:
LENGTH: 471 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-961-083-158

Query Match
Best Local Similarity 86.2%; Score 25; DB 4; Length 471;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EIIIMD 6
Db 222 DEVIDM 227

RESULT 15
US-09-019-942-3
Sequence 3, Application US/09019942
Patent No. 6033855
GENERAL INFORMATION:
APPLICANT: Berlin, John
TITLE OF INVENTION: GENES ENCODING CASPASE RECRUITMENT
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/019,942
FILING DATE: 06-FEB-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Melkielehn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07334/068001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 200 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-019-942-3

Query Match
Best Local Similarity 82.8%; Score 24; DB 3; Length 200;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Tue Mar 19 11:00:04 2002

us-09-544-665-3.rtf

Page 7

OY 1 ERTMD 6
|||
Db 115 ERTMD 120

Search completed: March 18, 2002, 09:49:41
Job time: 225 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:50:30 ; Search time 41.98 Seconds
(without alignments)
10.887 Million cell updates/sec

US-09-544-665-3

Title: US-09-544-665-3
Perfect score: 29
Sequence: 1 ERIIMD 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR_68:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	100.0	197	2	SO1240
2	29	100.0	197	2	H85907
3	29	100.0	284	2	B84616
4	29	100.0	400	2	JC4265
5	29	100.0	402	1	IRHUP1
6	29	100.0	402	1	SO6745
7	29	100.0	525	2	EF0125
8	27	93.1	301	2	SS1439
9	27	93.1	574	2	T27100
10	26	89.7	120	2	SS3052
11	26	89.7	240	2	JM0407
12	26	89.7	274	2	D71052
13	26	89.7	274	2	F75181
14	26	89.7	284	2	C72320
15	26	89.7	288	2	A70371
16	26	89.7	292	2	FE4486
17	26	89.7	346	2	DB4833
18	26	89.7	424	2	T10985
19	26	89.7	508	2	T64032
20	26	89.7	662	2	T20343
21	26	89.7	739	2	G36740
22	26	89.7	740	1	B65136
23	26	89.7	740	1	H86005
24	26	89.7	762	1	A64420
25	25	86.2	120	2	C72425
26	25	86.2	139	2	I40604
27	25	86.2	237	2	F86789
28	25	86.2	243	2	I51746
29	25	86.2	245	2	C72382

30	25	86.2	322	2	T24525	hypothetical prote
31	25	86.2	331	2	A75480	lipotic acid syntha
32	25	86.2	361	2	D75039	hypothetical prote
33	25	86.2	389	2	B69096	corrinoid/iron-sul
34	25	86.2	411	2	G86768	ATP dependent Clp
35	25	86.2	414	2	F64409	hypothetical prote
36	25	86.2	441	2	G64492	hypothetical prote
37	25	86.2	471	2	T21972	hypothetical prote
38	25	86.2	494	2	S67314	regulatory protein
39	25	86.2	540	2	JC7381	DNA-directed DNA p
40	25	86.2	591	2	UQ2036	hypothetical ORF p
41	25	86.2	684	2	D71683	proteolase II (ptr
42	25	86.2	719	2	T27977	lin-15A protein -
43	25	86.2	820	2	T41978	helicase - human h
44	25	86.2	1047	2	T25782	hypothetical prote
45	25	86.2	1084	2	A28555	middle wall protel

ALIGNMENTS

RESULT 1
SO1240
heat shock protein grpE (heat shock protein b25.3) (hsp24) - Escherichia coli
N:Alternate names: grpE protein
C:Species: Escherichia coli
C:Date: 28-Aug-1989 #sequence_revision 28-Aug-1989 #text_change 24-Sep-1999
R:Liipinska, B.; King, J.; Ang, D.; Georgopoulos, C.
Nucleic Acids Res. 16: 7545-7562, 1988
A:Title: Sequence analysis and transcriptional regulation of the Escherichia coli grp
A:Reference number: SO1240; KIID:8519942
A:Accession: SO1240
A:Molecule type: DNA
A:Residues: 1-197 <LRP>
A:Cross-references: EMBL:X07863; NID:941618; PIDN:CAA30711.1; PID:941619
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; KIID:97426617
A:Accession: A65040
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-197 <BLAT>
A:Cross-references: GB:AE000347; GB:U00096; NID:92367142; PIDN:AACT5663.1; PID:917889
A:Experimental source: strain K-12, substrain MG1655
A:Genetics:
A:Gene: grpE
A:Superfamily: heat shock protein grpE
F:2-197/Product: heat shock protein B25.3 #status predicted <MAT>

Query Match 100.0%; Score 29; DB 2; Length 197;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERIIMD 6
Db 15 ERIIMD 20

RESULT 2
H85907
hypothetical protein grpE [imported] - Escherichia coli (strain O157:H7)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Mar-2001
C:Accession: H85907
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimmlanta, E.; Potamoussis, K.; Apoda
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; KIID:21074935; PMID:11206551

A:Accession: H85907
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-197 <STO>
A:Cross-references: GB:AE005174; NID:912517038; PIDN:AA657724.1; GSPDB:GN00145; UMGF:239
A:Experimental source: strain O157:H7, substrain EDL933
A:Genetics:
A:Gene: grpe
C:Superfamily: heat shock protein grpe

Query Match 100.0%; Score 29; DB 2; Length 197;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ERIIMD 6
|||||
DB 15 ERIIMD 20

RESULT 3
E84616
Probable bHLH transcription factor [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: E84616
R:Lin. X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M. Koo, H.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487
A:Accession: E84616
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-284 <STO>
A:Cross-references: GB:AE002093; NID:93738090; PIDN:AA63587.1; GSPDB:GN00139
A:Genetics:
A:Gene: At2g22760
A:Map position: 2

Query Match 100.0%; Score 29; DB 2; Length 284;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ERIIMD 6
|||||
DB 90 ERIIMD 95

RESULT 4
JC4265
Plasminogen activator inhibitor type 1 precursor - American mink
C:Species: Mustela vison (American mink)
C>Date: 19-Oct-1995 #sequence_revision 08-Feb-1996 #text_change 20-Jun-2000
C:Accession: JC4265
R:Chuang, T.H.; Hamilton, R.T.; Nilsen-Hamilton, M.
Gene 162, 303-308, 1995
A:Title: Cloning of the mink plasminogen activator inhibitor type-1 messenger RNA: An mR
A:Reference number: JC4265; MUID:96032362
A:Accession: JC4265
A:Molecule type: mRNA
A:Residues: 1-400 <CHU>
A:Cross-references: EMBL:X58541; NID:91164923; PIDN:CAA41433.1; PID:91164924
A:Experimental source: lung CCL64 epithelial cells
A:Comment: This protein controls the activities of the plasminogen activators and plasm
A:Genetics:
A:Gene: pai-1
C:Superfamily: antithrombin III
C:Keywords: glycoprotein; plasminogen activator; serine proteinase inhibitor
F:1/21/Domain: signal sequence #status predicted <SIG>
F:22-400/Product: plasminogen activator inhibitor type 1 #status predicted <MAT>

F:230,286,350/Binding site: carbohydrate (asn) (covalent) #status predicted
F:367/Inhibitory site: Arg (plasminogen activator) #status predicted

Query Match 100.0%; Score 29; DB 2; Length 400;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ERIIMD 6
|||||
DB 371 ERIIMD 376

RESULT 5
ITRHP1
plasminogen activator inhibitor 1 precursor [validated] - human
N:Alternate names: plasminogen activator inhibitor, endothelial
C:Species: Homo sapiens (man)
C>Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 08-Dec-2000
C:Accession: A28107; S02551; A26996; I59126; J50397; A25693; A26146; A29100; A25895;
R:Bohma, P.J.; van den Berg, E.A.; Koolstra, T.; Stenlund, D.R.; Slighem, J.L.
J. Biol. Chem. 265, 9129-9141, 1988
A:Title: Human plasminogen activator inhibitor-1 gene. Promoter and structural gene n
A:Reference number: A28107; MUID:88243790
A:Accession: A28107
A:Molecule type: DNA
A:Residues: 1-402 <BO>
A:Cross-references: GB:U03764; NID:9189564; PIDN:AA60007.1; PID:9386996
R:Salzberg, L.; Lawrence, D.; NY, T.
Eur J. Biochem. 176, 609-616, 1988
A:Title: The organization of the human-plasminogen-activator-inhibitor-1 gene. Implic
A:Reference number: S02551; MUID:89005111
A:Accession: S02551
A:Molecule type: DNA
A:Residues: 1-14, T, 16-402 <STR>
A:Cross-references: EMBL:X1338; NID:935244; PIDN:CAA31722.1; PID:935245
A:Note: the complete translation is not annotated in Genbank entries HSPAI12
R:Loskutoff, D.J.; Linders, M.; Kelfer, J.; Veerman, H.; van Heerikhuizen, H.; Pannek
Biochemistry 26, 3763-3768, 1987
A:Title: Structure of the human plasminogen activator inhibitor 1 gene: nonrandom dis
A:Reference number: A26996; MUID:8800586
A:Accession: A26996
A:Molecule type: DNA
A:Residues: 1-402 <LOS>
A:Cross-references: GB:A22221; GB:M17121; NID:9189576; PIDN:AA60009.1; PID:9189578
A:Note: the sequence in Genbank entry H0MPA1B2, release 109.0, (PID:9189578) has the
R:van Zonneveld, A.J.; Curriden, S.A.; Loskutoff, D.J.
Proc. Natl. Acad. Sci. U.S.A. 85, 5525-5529, 1988
A:Title: Type 1 plasminogen activator inhibitor gene: functional analysis and glucoco
A:Reference number: I59126; MUID:88289754
A:Accession: I59126
A>Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-9 <ON>
A:Cross-references: GB:J03836; NID:9189579; PIDN:AA60010.1; PID:9189580
R:Polio, M.; Ginsburg, D.
Gene 88, 447-453, 1989
A:Title: Structure and expression of the human gene encoding plasminogen activator in
A:Reference number: J50397; MUID:90128289
A:Accession: J50397
A:Molecule type: DNA
A:Residues: 85-86; 88-93; 166-171; 231-236; 298-302; 331-336; 360-365; 388-393 <FOU>
A:Cross-references: GB:M33136; NID:9189543
A:Note: sequences of the Intron/exon boundaries are shown
R:Pannekoek, H.; Veerman, H.; Lamberts, H.; Diergaarde, P.; Verwel, C.L.; van Zonneve
EMBO J. 5, 2539-2544, 1986
A:Title: Endothelial plasminogen activator inhibitor (PAI): a new member of the serpi
A:Reference number: A91052; MUID:87053819
A:Accession: A25693
A:Molecule type: mRNA
A:Residues: 1-402 <PAN>
A:Cross-references: GB:X04429; NID:935271; PIDN:CAA28025.1; PID:935272
R:Ginsburg, D.; Zehnb, R.; Yang, A.Y.; Rafferty, U.M.; Andreasen, P.A.; Nielsen, L.;

J. Clin. Invest. 78, 1673-1680, 1986
 A:Title: cDNA cloning of human plasminogen activator-inhibitor from endothelial cells.
 A:Reference number: A92766; MUID:87058123
 A:Accession: A26146
 A:Molecule type: mRNA
 A:Residues: 1-402 <GIN>
 A:Cross-references: GB:M16006; NID:g189541; PIDN:AAA60003.1; PID:g189542
 R:Min, T.C.; Kretzmer, K.K.
 FEBS Lett. 210, 11-16, 1987
 A:Title: cDNA cloning and expression in E. coli of a plasminogen activator inhibitor (PAI)
 A:Reference number: A29100; MUID:87105925
 A:Accession: A29100
 A:Molecule type: mRNA
 A:Residues: 17-402 <MUN>
 A:Cross-references: GB:X04744; NID:g35275; PIDN:CAA28444.1; PID:g755747
 A:Note: part of this sequence, including the amino end of the mature protein, was confid
 R:W, T.; Sawley, M.; Lawrence, D.; Millan, J.L.; Loskutoff, D.J.
 Proc. Natl. Acad. Sci. U.S.A. 83, 6776-6780, 1986
 A:Title: Cloning and sequence of a cDNA coding for the human beta-migrating endothelial
 A:Reference number: A25895; MUID:86313660
 A:Accession: A25895
 A:Molecule type: mRNA
 A:Residues: 20-402 <NT>
 A:Cross-references: GB:M14083; NID:g189566; PIDN:AAA60008.1; PID:g386997
 R:Andreasen, P.A.; Riccio, A.; Weiland, K.G.; Douglas, R.; Sartorio, R.; Nielsen, L.S.
 FEBS Lett. 209, 213-218, 1986
 A:Title: Plasminogen activator inhibitor type-1: reactive center and amino-terminal helix
 A:Reference number: A91371; MUID:87080762
 A:Accession: A25651
 A:Molecule type: mRNA
 A:Residues: 1-14, 16-47 <AND1>
 A:Cross-references: GB:X04729; NID:g35263; PIDN:CAA28438.1; PID:g35264
 A:Accession: B25651
 A:Molecule type: mRNA
 A:Residues: 364-402 <AND2>
 A:Cross-references: GB:X04731; NID:g35260; PIDN:CAA28442.1; PID:g35261
 R:Lang, W.E.; Abersold, R.; Jong, A.; Rideout, W.; Bergman, B.L.; Baker, J.
 Thromb. Haemost. 61, 517-521, 1989
 A:Title: Isolation of multiple types of plasminogen activator inhibitors from vascular
 A:Reference number: A60436; MUID:90020174
 A:Accession: A60436
 A:Molecule type: protein
 A:Residues: 225-235 <LUD>
 R:Kjoller, L.; Mathiesen, P.M.; Sottrup-Jensen, L.; Justesen, J.; Rodenbury, K.W.; Andre
 Eur. J. Biochem. 241, 38-46, 1996
 A:Title: Conformational changes of the reactive-centre loop and beta-strand 5A accompany
 A:Reference number: S74133; MUID:97054589
 A:Accession: S74133
 A:Molecule type: protein
 A:Residues: 22-30, 370-376 <KJO>
 R:Stromqvist, M.; Karlsson, K.E.; Bjorkqvist, P.; Andersson, J.O.; Bystrom, M.; Hansson
 Biochim. Biophys. Acta 1295, 103-109, 1996
 A:Title: Characterization of the complex of plasminogen activator inhibitor type 1 with
 A:Reference number: S70346; MUID:96283799
 A:Accession: S70346
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 370-375 <STO>
 C:Comment: This inhibitor acts as "bait" for tissue plasminogen activator (see PIR:UKHUT
 fibrinolysis. High concentrations of this protein have been associated with human throm
 C:Comment: Three types of PAI have been identified (see also PIR:A32853 and PIR:A39339).
 C:Comment: Vascular endothelial cells may be the primary site of synthesis of plasma PAI
 C:Comment: Glycosylation is not required for inhibitory activity.
 C:Genetics:
 A:Gene: GDB:PAI1; PLANH1
 A:Cross-references: GDB:120297; OMIM:173360
 A:Map position: 7q21.3-7q22
 A:Insertions: 91/1; 169/1; 234/1; 300/2; 334/1; 363/1; 391/1
 C:Superfamily: antithrombin III
 C:Keywords: glycoprotein; serine proteinase inhibitor
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-402/Product: plasminogen activator inhibitor-1 #status experimental <MAT>
 F:232,288,352/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:369/Inhibitory site: Arg (plasminogen activator) #status predicted

Query Match 100.0%; Score 29; DB 1; Length 402;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIIIMD 6
 DB 373 EIIIMD 378

RESULT 6
 S06745

Plasminogen activator inhibitor-1 precursor - bovine
 M:Alternate names: endothelial cell plasminogen activator inhibitor; PAI-1
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 28-Feb-1990 #sequence revision 22-Apr-1995 #text_change 18-Jun-1999
 C:Accession: S06745; A35855; S01324; S10906
 R:Minuro, J.; Sawley, M.; Hattori, M.; Loskutoff, D.J.
 Nucleic Acids Res. 17, 8872-1989
 A:Title: cDNA for bovine type 1 plasminogen activator inhibitor (PAI-1).
 A:Reference number: S06745; MUID:9067867
 A:Accession: S06745
 A:Molecule type: mRNA
 A:Residues: 1-402 <MUN>
 A:Cross-references: EMBL:X16383; NID:g600; PIDN:CAA34419.1; PID:g601
 R:Pepper, M.S.; Bellin, D.; Montesano, R.; Orci, L.; Vassalli, J.D.
 J. Cell Biol. 111, 743-755, 1990
 A:Title: Transforming growth factor-beta 1 modulates basic fibroblast growth factor-1
 A:Reference number: A35855; MUID:90338128
 A:Accession: A35855
 A:Molecule type: mRNA
 A:Residues: 153-235 <PEP>
 A:Cross-references: EMBL:X52906; NID:g598; PIDN:CAA37094.1; PID:g930005
 R:Kataagiri, K.; Okada, K.; Hattori, H.; Yano, M.
 Eur. J. Biochem. 176, 81-87, 1988
 A:Title: Bovine endothelial cell plasminogen activator inhibitor. Purification and he
 A:Reference number: S01324; MUID:88329072
 A:Accession: S01324
 A:Molecule type: protein
 A:Residues: 24-49, 51-63 <KAT>
 C:Comment: Three types of PAI have been identified. PAI-1 is an acid-stable glycoprot
 C:Comment: The primary site of synthesis of plasma PAI.
 C:Comment: This inhibitor acts as "bait" for tissue plasminogen activator, urokinase,
 yls.
 C:Comment: Glycosylation is not a prerequisite for inhibitory activity.
 C:Superfamily: antithrombin III
 C:Keywords: glycoprotein; serine proteinase inhibitor
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-402/Product: plasminogen activator inhibitor-1 #status predicted <MAT>
 F:232,288,352/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:369/Inhibitory site: Arg (plasminogen activator) #status predicted

Query Match 100.0%; Score 29; DB 1; Length 402;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIIIMD 6
 DB 373 EIIIMD 378

RESULT 7

hypothetical protein BB0205 - Lyme disease spirochete
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)
 C:Date: 13-Feb-1998 #sequence revision 13-Feb-1998 #text_change 28-Jul-2000
 C:Accession: E70125
 R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; W
 son, D.; Peterson, J.; Kariyase, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; V
 Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.

Nature 390, 580-586, 1997
 A:Authors: Smith, H.O.; Venter, J.C.
 A:Title: Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*.
 A:Reference number: A70100; MUID:98065943
 A:Accession: E70125
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-525 <RLF>
 A:Cross-references: GB:AE001131; GB:AE000783; NID:q2688098; PIDN:AAC66601.1; PID:q268801
 A:Experimental source: strain B31
 C:Superfamily: *Borrelia burgdorferi* hypothetical protein BB0205

Query Match 100.0%; Score 29; DB 2; Length 525;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 EELIMD 6
 Db 461 EELIMD 466

RESULT 8
 S51439
 hypothetical protein YLR307w - yeast (*Saccharomyces cerevisiae*)
 N:Alternate names: hypothetical protein L2142.2
 C:Species: *Saccharomyces cerevisiae*
 C:Date: 23-Feb-1995 #sequence_revision 12-May-1995 #text_change 05-Nov-1999
 C:Accession: S51439
 R:Pauley, A.
 submitted to the EMBL Data Library, November 1994
 A:Description: The sequence of *S. cerevisiae* cosmid L2142.
 A:Reference number: S51437
 A:Accession: S51439
 A:Molecule type: DNA
 A:Residues: 1-301 <PAU>
 A:Cross-references: EMBL:U17247; NID:9577216; PIDN:AAB67356.1; PID:9577218; GSPDB:GN0001
 C:Genetics:
 A:Gene: MIPS:YLR307w
 A:Map position: 12R
 C:Superfamily: nodB homology
 F:113-264/Domain: nodB homology <NODB>

Query Match 93.1%; Score 27; DB 2; Length 301;
 Best Local Similarity 83.3%; Pred. No. 67;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 EELIMD 6
 Db 243 EELIMD 248

RESULT 9
 T27100
 hypothetical protein Y51H1A.6 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
 C:Accession: T27100
 R:Smyle, R.
 submitted to the EMBL Data Library, October 1998
 A:Reference number: Z20309
 A:Accession: T27100
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-574 <RTL>
 A:Cross-references: EMBL:AL022644; PIDN:CAA21671.1; GSPDB:GN00020; CESP:Y51H1A.6
 C:Experimental source: clone Y51H1A
 C:Genetics:
 A:Gene: CESP:Y51H1A.6
 A:Map position: 2
 A:Accession: J2211
 C:Superfamily: *Caenorhabditis elegans* hypothetical protein Y51H1A.6

Query Match 93.1%; Score 27; DB 2; Length 574;
 Best Local Similarity 66.7%; Pred. No. 1,4e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 EELIMD 6
 Db 130 EELIMD 135

RESULT 10
 S63052
 cytochrome b5 - yeast (*Saccharomyces cerevisiae*)
 N:Alternate names: protein N1949; protein YNL111c
 C:Species: *Saccharomyces cerevisiae*
 C:Date: 27-Apr-1996 #sequence_revision 03-May-1996 #text_change 20-Jun-2000
 C:Accession: S63052; S47919; S67347
 R:De Antoni, A.; D'Angelo, M.; Dal Pero, F.; Sartorello, F.; Pandolfo, D.; Lanfranchi
 submitted to the Protein Sequence Database, April 1996
 A:Reference number: S63047
 A:Accession: S63052
 A:Molecule type: DNA
 A:Residues: 1-120 <DEA>
 A:Cross-references: EMBL:Z71387; NID:q1302031; PIDN:CAA95990.1; PID:q1302032; MIPS:YN
 R:Tuan, G.; Epinat, J.C.; Rougeulle, C.; Cullin, C.; Pompon, D.
 Gene 149, 123-127, 1994
 A:Title: Cloning and characterization of a yeast cytochrome b(5)-encoding gene which
 A:Reference number: S47919; MUID:95047457
 A:Accession: S47919
 A:Molecule type: DNA
 A:Residues: 1-16, 'Q', 18-120 <TRU>
 A:Cross-references: EMBL:U22484; NID:q431761; PIDN:AAA67468.1; PID:q431762
 R:d'Antoni, A.; d'Angelo, M.; dal Pero, F.; Sartorello, F.; Pandolfo, D.; Lanfranchi,
 submitted to the EMBL Data Library, February 1996
 A:Description: The DNA sequence of cosmid 14-13b from chromosome XIV of *Saccharomyces*
 A:Reference number: S67327
 A:Accession: S67347
 A:Molecule type: DNA
 A:Residues: 1-120 <DAN>
 A:Cross-references: EMBL:Z69382; NID:q1183941; PIDN:CAA93396.1; PID:q1183962
 C:Genetics:
 A:Gene: SGD:CYB5
 A:Cross-references: SGD:S0005055; MIPS:YNL111c
 A:Map position: 14L
 C:Function:
 A:Description: electron transfer
 C:Superfamily: cytochrome b5; cytochrome b5 core homology
 C:Keywords: electron transfer; heme; iron; metalloprotein; transmembrane protein
 F:2-77/Domain: cytochrome b5 core homology <CB5>
 F:102-118/Domain: transmembrane #status predicted <TM>
 F:37,61/Binding site: heme iron (His) (axial ligands) #status predicted

Query Match 89.7%; Score 26; DB 2; Length 120;
 Best Local Similarity 83.3%; Pred. No. 41;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 EELIMD 6
 Db 41 EELIMD 46

RESULT 11
 JN0407
 polyhedrin - *Galleria mellonella* nuclear polyhedrosis virus (tentative sequence) (fra
 N:Alternate names: inclusion body protein
 C:Species: *Galleria mellonella* nuclear polyhedrosis virus
 C:Date: 15-Jan-1993 #sequence_revision 15-Jan-1993 #text_change 31-Mar-2000
 C:Accession: JN0407
 R:Gusak, N.M.; Kozlov, E.A.; Ovander, M.N.; Serebriyany, S.B.
 Bioorg. Khim. 7, 996-1007, 1981

A:Title: tryptic peptides of inclusion body protein of nuclear polyhedrosis virus of *Ga*
A:Reference number: JN0407
A:Accession: JN0407
A:Molecule type: protein
A:Residues: 1-240 <GUS>
A:Note: 18-Arg, 87-Ala, 87-Gly, 126-Met and 139-Glu were also found
C:Superfamily: nuclear polyhedrosis virus polyhedrin
C:Keywords: polyhedrin

Query Match 89.7%; Score 26; DB 2; Length 240;
Best Local Similarity 50.0%; Pred. No. 90;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIIMD 6
:|||||
Db 112 ZEYVMD 117

RESULT 12

D71032
probable autoantigen like protein - *Pyrococcus horikoshii*
C:Species: *Pyrococcus horikoshii*
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000
C:Accession: D71032
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Seki,
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kushiida, N.; Oguchi,
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A:Reference number: A71000; MUID:98344137
A:Accession: D71032
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-274 <KAM>
A:Cross-references: GB:AP000006; NID:93236133; PIDN:BAA30660.1; PID:93257977
A:Experimental source: strain OR3
A:Note: this accession replaces an interim accession for a sequence replaced by Genbank
C:Genetics:
A:Gene: PH1548
C:Superfamily: conserved hypothetical protein MTH682

Query Match 89.7%; Score 26; DB 2; Length 274;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIIMD 6
:|||||
Db 213 EELVMD 218

RESULT 13

F75181
POLYribonucleotide nucleotidyltransferase related protein PAB0421 - *Pyrococcus abyssi* (S
C:Species: *Pyrococcus abyssi*
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 21-Jul-2000
C:Accession: F75181
R:anonymous, Genoscope
Submitted to the EMBL Data Library, July 1999
A:Description: *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome stru
A:Reference number: A75001
A:Accession: F75181
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-274 <KAM>
A:Cross-references: GB:AJ248284; GB:AI096836; NID:95457730; PIDN:CABA9533.1; PID:9545804
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB0421
C:Superfamily: conserved hypothetical protein MTH682

Query Match 89.7%; Score 26; DB 2; Length 274;

Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIIMD 6
:|||||
Db 213 EELVMD 218

RESULT 14

C72320
conserved hypothetical protein - *Thermotoga maritima* (strain MSB8)
C:Species: *Thermotoga maritima*
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: C72320
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hic
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,
C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
A:Reference number: A72200; MUID:99287316
A:Accession: C72320
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-284 <ARN>
A:Cross-references: GB:AE001755; GB:AE000512; NID:94981432; PIDN:AA035997.1; PID:9498
A:Experimental source: strain MSB
C:Genetics:
A:Gene: TM0916
C:Superfamily: Aeropyrum pernix hypothetical protein APE2212

Query Match 89.7%; Score 26; DB 2; Length 284;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIIMD 6
:|||||
Db 139 DEIIMD 144

RESULT 15

A70371
conserved hypothetical protein aq_814 - *Aquifex aeolicus*
C:Species: *Aquifex aeolicus*
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 20-Jun-2000
C:Accession: A70371
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.
A:Reference number: A70300; MUID:98196666
A:Accession: A70371
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-288 <AOE>
A:Cross-references: GB:AE000708; NID:92983356; PIDN:AAC06959.1; PID:92983372; GB:AE0C
A:Experimental source: strain VFS
C:Genetics:
A:Gene: aq_814
C:Superfamily: Aeropyrum pernix hypothetical protein APE2212

Query Match 89.7%; Score 26; DB 2; Length 288;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIIMD 6
:|||||
Db 142 DEIIMD 147

Search completed: March 18, 2002, 09:50:31

Tue Mar 19 11:00:04 2002

us-09-544-665-3.rpr

Page 6

Job time: 240 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:52:25 ; Search time 24.53 Seconds

(without alignments)
8.968 Million cell updates/sec

Title: US-09-544-665-3

Perfect score: 29

Sequence: 1 EELMD 6

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	100.0	197	1 GRPE_ECOLI	P09372 escherichia
2	29	100.0	400	1 PAII_MOSVI	P50449 mustela vis
3	29	100.0	402	1 PAII_BOVIN	P13309 bos taurus
4	29	100.0	402	1 PAII_HUMAN	P05121 homo sapien
5	29	100.0	402	1 PAII_PIG	P79335 sus scrofa
6	27	93.1	301	1 CDAA_YEAST	O06702 saccharomyc
7	26	89.7	120	1 CYB5_YEAST	P40312 saccharomyc
8	26	89.7	147	1 RPB6_ASFB7	P42484 african swi
9	26	89.7	292	1 YB95_METJA	O58890 methanococc
10	26	89.7	508	1 YG28_HAEIN	P44224 haemophilus
11	26	89.7	662	1 YLNU_CAEEL	O09512 caenorhabdi
12	26	89.7	762	1 Y961_METJA	O58371 methanococc
13	26	89.7	773	1 YHGF_ECOLI	P46837 escherichia
14	25	86.2	139	1 CDRB_CLOAB	O45811 clostridium
15	25	86.2	139	1 RPB6_ASFM2	O08399 african swi
16	25	86.2	322	1 YJ96_CAEEL	P49048 caenorhabdi
17	25	86.2	331	1 LTPA_DEIRA	O91944 delnecoccus
18	25	86.2	411	1 CLPX_LACLA	O93964 lactococcus
19	25	86.2	414	1 Y878_METJA	O58388 methanococc
20	25	86.2	820	1 H87L_HSV73	P52357 human herpe
21	25	86.2	1053	1 SLPM_BACBR	P06546 bacillus br
22	25	86.2	2238	1 RRP1_BUNYM	P20470 bunyamwera
23	24	82.8	123	1 CS9A_MOUSE	O55186 mus musculu
24	24	82.8	150	1 UREE_STRSL	O55055 streptococc
25	24	82.8	170	1 RL5_THEAC	O56231 thermoplasm
26	24	82.8	245	1 REVY_YEAST	P38927 saccharomyc
27	24	82.8	253	1 KANU_BACSP	P05058 bacillus sp
28	24	82.8	253	1 KANU_STRAU	P05057 staphylococ
29	24	82.8	254	1 PDHR_ECOLI	P06957 escherichia
30	24	82.8	290	1 ARGB_ARCFU	O28988 archaeeoglob
31	24	82.8	293	1 NIFH_AZOBH	P17303 azospirillum
32	24	82.8	293	1 RRP1_HUMAN	O13868 homo sapien
33	24	82.8	295	1 NIF1_RHOCA	P08718 rhodobacter

34	24	82.8	353	1 NUSG_THEMA	P29397 thermotoga
35	24	82.8	353	1 REON_HUMAN	O92782 homo sapien
36	24	82.8	373	1 YCO8_KLEPN	O48454 klebsiella
37	24	82.8	387	1 REON_MOUSE	O99x66 mus musculu
38	24	82.8	391	1 TAL_SYNY3	P72797 synecocyst
39	24	82.8	397	1 REON_RAT	P56163 rattus norv
40	24	82.8	429	1 URAA_ECOLI	P33780 escherichia
41	24	82.8	431	1 AROA_AQUAE	O67494 aquifex neo
42	24	82.8	455	1 YNE2_CAEEL	P30641 caenorhabdi
43	24	82.8	459	1 UUP2_HAEIN	P43167 haemophilus
44	24	82.8	508	1 RPA2_THEAC	O03586 thermoplasm
45	24	82.8	526	1 MITF_HUMAN	O75030 homo sapien

ALIGNMENTS

RESULT 1
GRPE_ECOLI STANDARD: PRT: 197 AA.
ID GRPE_ECOLI
AC P09372:
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE GRPE PROTEIN (HSP-70 COFACTOR) (HEAT SHOCK PROTEIN B25.3) (HSP24).
GN GRPE OR B2614.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID:562;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-B178;
RX MEDLINE-88319942; PubMed-3045760;
RA Lipinska B., King J., Ang D., Georgopoulos C.;
RT "Sequence analysis and transcriptional regulation of the Escherichia coli grpe gene, encoding a heat shock protein.";
RL Nucleic Acids Res. 16:7545-7562(1988).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE-97426617; PubMed-9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN (3)
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE-97349980; PubMed-9205837;
RA Yamamoto Y., Alta H., Baba T., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nishimoto H.,
RA Oshima T., Oyama S., Saito N., Sempel G., Setoh Y., Sivasubram S.,
RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA Yamagata S., Horiuchi T.;
RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
- K12 genome corresponding to 50.0-68.8 min on the linkage map and
analysis of its sequence features.";
RL DNA Res. 4:91-113(1997).
RN (4)
RP POSSIBLE FUNCTION.
RX MEDLINE-91187894; PubMed-1826368;
RA "Liberek K., Marzalek J., Ang D., Georgopoulos C., Zyllicz M.;
RT "Escherichia coli Dna and Grpe heat shock proteins jointly stimulate
ATPase activity of DnaK.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:2874-2878(1991).
RN (5)
RP FUNCTION, AND MUTAGENESIS.
RX MEDLINE-97045089; PubMed-8890154;
RA Wu B., Mawczynow A., Zyllicz M., Georgopoulos C.;

RT *Structure-function analysis of the Escherichia coli GrpE heat shock
 RT protein.";
 RL EMBL J. 15:4806-4816(1996).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
 RX MEDLINE-97258949; PubMed-9103205;
 RA Harrison C.J., Hayer-Hartl M., di Liberto M., Hartl F., Kuriyan J.;
 RT "Crystal structure of the nucleotide exchange factor GrpE bound to
 the ArpE domain of the molecular chaperone DnaK.";
 RL Science 276:431-435(1997).
 CC -1- FUNCTION: STIMULATES, JOINTLY WITH DnaJ, THE ATPASE ACTIVITY OF
 DnaK. HELPS TO RELEASE ADP FROM DnaK THUS ALLOWING DnaK TO RECYCLE
 MORE EFFICIENTLY.
 CC -1- INDUCTION: BY THE SIGMA(32) SUBUNIT OF RNA POLYMERASE.
 CC -1- SIMILARITY: BELONGS TO THE GRPE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X07863; CAA30711.1; -
 CC EMBL: AE000347; AAC75663.1; -
 CC EMBL: D90888; BAA16498.1; -
 CC PIR: S01240; S01240.
 CC PDB: 1DKG; 20-AUG-97.
 CC DR SWISS-2DPAGE; P09372; COLI.
 CC DR ECODBASE; B0253; 6TH EDITION.
 CC DR EcoGene; EGI0416; grpe.
 CC DR InterPro: IPR000740; Grpe.
 CC DR Pfam: PF01025; Grpe; 1.
 CC DR PRINTS: PR00773; GREPROTEIN.
 CC DR PROSITE: PS01071; GRPE; 1.
 CC DR Chapterone; Heat shock; 3D-structure; Complete proteome.
 CC KW SEQUENCE 197 AA; 21798 MW; CDC4CD9D08AD4BEF CRC64;
 SQ

Query Match 100.0%; Score 29; DB 1; Length 197;
 Best Local Similarity 100.0%; Pred. NO. 5.4; Indels 0; Gaps 0;
 Matches 6; Conservative 0; Mismatches 0;
 OY 1 EELIND 6
 DB 15 EELIND 20

RESULT 2
 PAI_MUSVI STANDARD; PRT; 400 AA.
 AC P50449;
 DT 01-OCT-1996 (rel. 34, Last sequence update)
 DT 01-OCT-1996 (rel. 40, Last annotation update)
 DE 20-AUG-2001 (rel. 40, Last annotation update)
 DE PLASMINOGEN ACTIVATOR INHIBITOR-1 PRECURSOR (PAI-1) (ENDOTHELIAL
 GN SERPINE1 OR PAI1 OR PLANH1 OR PAI-1).
 OS Mus musculus (American mink).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Mustelidae; Mustela.
 OX NCBI_Taxid=9667;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=lung;
 RX MEDLINE-96032362; PubMed-7557448;
 RA Chuang T.H., Hamilton R.T., Nilsen-Hamilton M.;
 RT "Cloning of the mink plasminogen activator inhibitor type-1 messenger
 RL RNA: an mRNA with a short half life.";
 CC Gene 162:303-308(1995).
 CC -1- FUNCTION: THIS INHIBITOR ACTS AS "BAIT" FOR TISSUE PLASMINOGEN
 CC ACTIVATOR, UROKINASE, AND PROTEIN C. ITS RAPID INTERACTION WITH

CC TPA MAY FUNCTION AS A MAJOR CONTROL POINT IN THE REGULATION OF
 CC FIBRINOLYSIS.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X58541; CAA41433.1; -
 CC DR HSSP; P01008; IATH.
 CC DR InterPro: IPR000215; Serpin.
 CC DR Pfam; PF00079; serpin; 1.
 CC DR SMART; SM00093; SERPIN; 1.
 CC DR PROSITE; PS00284; SERPIN; 1.
 CC KW Serpin; Serine protease inhibitor; Plasma; Plasminogen activation;
 CC Glycoprotein; Signal.
 CC FT SIGNAL 1 21 POTENTIAL.
 CC FT CHAIN 22 400 PLASMINOGEN ACTIVATOR INHIBITOR-1.
 CC FT ACT_SITE 367 368 REACTIVE BOND (BY SIMILARITY).
 CC FT CARBOHYD 230 230 N-LINKED (GLCNAC...) (POTENTIAL).
 CC FT CARBOHYD 286 286 N-LINKED (GLCNAC...) (POTENTIAL).
 CC FT CARBOHYD 350 350 N-LINKED (GLCNAC...) (POTENTIAL).
 CC FT SEQUENCE 400 AA; 45153 MW; DF45E0694DE28401 CRC64;
 SQ

Query Match 100.0%; Score 29; DB 1; Length 400;
 Best Local Similarity 100.0%; Pred. NO. 12;
 Matches 6; Conservative 0; Mismatches 0;
 OY 1 EELIND 6
 DB 371 EELIND 376

RESULT 3
 PAI_BOVIN STANDARD; PRT; 402 AA.
 AC P13909;
 DT 01-JAN-1990 (rel. 13, Created)
 DT 01-JAN-1990 (rel. 13, Last sequence update)
 DT 20-AUG-2001 (rel. 40, Last annotation update)
 DE PLASMINOGEN ACTIVATOR INHIBITOR-1 PRECURSOR (PAI-1) (ENDOTHELIAL
 GN SERPINE1 OR PAI1).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_Taxid=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE-9067867; PubMed-2587231;
 RX Minoro J., Sawley M., Hattori M., Loskutov D.J.;
 RT "cDNA for bovine type 1 plasminogen activator inhibitor (PAI-1).";
 RL Nucleic Acids Res. 17:8872-8872(1989).
 RN [2]
 RP SEQUENCE OF 24-63; PubMed-3262060;
 RX MEDLINE-8839072; PubMed-3262060;
 RA Katagiri K., Okada K., Hattori H., Yano M.;
 RT "Bovine endothelial cell plasminogen activator inhibitor.
 RL Purification and heat activation.";
 CC Eur. J. Biochem. 176:81-87(1988).
 RN [3]
 RP SEQUENCE OF 153-235 FROM N.A.
 RC TISSUE=adrenal cortex;
 RX MEDLINE-9038128; PubMed-1696269;
 RA Pepper M.S., Belin D., Montesano R., Orci L., Vassalli J.-D.;
 RT "Transforming growth factor-beta 1 modulates basic fibroblast growth
 factor-induced proteolytic and angiogenic properties of endothelial

RT cells in vitro. ;
 RL J. Cell Biol. 111:743-755(1990).
 CC -1- FUNCTION: THIS INHIBITOR ACTS AS "BAIT" FOR TISSUE PLASMINOGEN
 CC ACTIVATOR, UROKINASE, AND PROTEIN C. ITS RAPID INTERACTION WITH
 CC TPA MAY FUNCTION AS A MAJOR CONTROL POINT IN THE REGULATION OF
 CC FIBRINOLYSIS.
 CC -1- MISCELLANEOUS: PAI1 IS INACTIVATED BY PROTEOLYTIC ATTACK OF THE
 CC UROKINASE-TYPE (U-PA) AND THE TISSUE-TYPE (TPA), CLEAVING THE
 CC 369(R)-370(M) BOND.
 CC -1- MISCELLANEOUS: VASCULAR ENDOTHELIAL CELLS MAY BE THE PRIMARY SITE
 CC OF SYNTHESIS OF PLASMA PAI.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC
 DR EMBL: X16383; CAA34419.1; ;
 DR EMBL: X52906; CAA37094.1; ;
 DR PIR: S01324; S01324.
 DR PIR: S06745; S06745.
 DR HSSP: P01008; 1ATH.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; Serpin; 1.
 DR SMART: SM00093; Serpin; 1.
 DR PROSITE: PS00284; SERPIN; 1.
 KW Serpin; Serine protease inhibitor; Plasma; Plasminogen activation;
 KW Glycoprotein; Signal.
 FT SIGNAL 1 23
 FT CHAIN 1 24
 FT ACT_SITE 369 370 REACTIVE BOND.
 FT CARBOHYD 232 232 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 288 288 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 50 50 S -> L (IN REF. 2).
 SO SEQUENCE 402 AA; 45371 MW; 905361733C7D130 CRC64;

 QY 1 EETIMD 6
 DB 373 EETIMD 378

 Query Match 100.0%; Score 29; DB 1; Length 402;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 RESULT 4
 PAI1_HUMAN STANDARD; PRT; 402 AA.
 ID PAI1_HUMAN
 AC P05121;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE PLASMINOGEN ACTIVATOR INHIBITOR-1 PRECURSOR (PAI-1) (ENDOTHELIAL
 DE PLASMINOGEN ACTIVATOR INHIBITOR) (PAI).
 GN SERPIN1 OR PAI1 OR PLANH1.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE-87053819; PubMed-2430793;
 RA Pannekoek H., Veerman H., Lambers H., Diergaarde P., Verweij C.L.,
 RA van Zonneveld A.-J., van Mourik J.A.;
 RT "Endothelial plasminogen activator inhibitor (PAI): a new member of
 RT the Serpin gene family.";
 RL EMBO J. 5:2539-2544(1986).

RN (2)
 RP SEQUENCE FROM N.A.
 RX MEDLINE-88000586; PubMed-2820474;
 RA Loskutoff D.J., Linders M., Keijser J., Veerman H.,
 RA van Heerikhuizen H., Pannekoek H.;
 RT "Structure of the human plasminogen activator inhibitor 1 gene:
 RT nonrandom distribution of introns.";
 RL Biochemistry 26:3763-3768(1987).
 RN (3)
 RP SEQUENCE FROM N.A.
 RX MEDLINE-87058123; PubMed-3097076;
 RA Ginsburg D., Zeneb R., Yang A.Y., Rafferty U.M., Andreasen P.A.,
 RA Nielsen L., Dano K., Lebo R.V., Gelehrter T.D.;
 RT "cDNA cloning of human plasminogen activator-inhibitor from
 RT endothelial cells.";
 RL J. Clin. Invest. 78:1673-1680(1986).
 RN (4)
 RP SEQUENCE FROM N.A.
 RX MEDLINE-90128289; PubMed-2612914;
 RA Folio M., Ginsburg D.;
 RT "Structure and expression of the human gene encoding plasminogen
 RT activator inhibitor, PAI-1.";
 RL Gene 84:447-453(1989).
 RN (5)
 RP SEQUENCE FROM N.A.
 RX MEDLINE-89005111; PubMed-3262512;
 RA Strandberg L., Lawrence D., Ny T.;
 RT "The organization of the human-plasminogen-activator-inhibitor-1
 RT gene. Implications on the evolution of the serine-protease inhibitor
 RT family.";
 RL Eur. J. Biochem. 176:609-616(1988).
 RN (6)
 RP SEQUENCE FROM N.A.
 RX MEDLINE-88243790; PubMed-3132455;
 RA Bosma P.J., van den Berg E.A., Koolstra T., Slemeniak D.R.,
 RA Slightom J.L.;
 RT "Human plasminogen activator inhibitor-1 gene. Promoter and
 RT structural gene nucleotide sequences.";
 RL J. Biol. Chem. 263:9129-9141(1988).
 RN (7)
 RP SEQUENCE FROM N.A.
 RA Pannekoek H.;
 RT Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
 RN (8)
 RP SEQUENCE FROM N.A.
 RA Cordes M., Doela D.;
 RT Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN (9)
 RP SEQUENCE OF 20-402 FROM N.A.
 RX MEDLINE-86313660; PubMed-3092219;
 NY T., Savdey M., Lawrence D., Millan J.L., Loskutoff D.J.;
 RT "Cloning and sequence of a cDNA coding for the human beta-migrating
 RT endothelial-cell-type plasminogen activator inhibitor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:6776-6780(1986).
 RN (10)
 RP SEQUENCE OF 1-47 AND 364-402 FROM N.A.
 RX MEDLINE-87080765; PubMed-3025016;
 RA Andreasen P.A., Riccio A., Wellinder K.G., Douglas R., Sartorio R.,
 RA Nielsen L.S., Oppenheimer C., Blasi F., Danoe K.;
 RT "Plasminogen activator inhibitor type-1: reactive center and amino-
 RT terminal heterogeneity determined by protein and cDNA sequencing.";
 RL FEBS Lett. 209:213-218(1986).
 RN (11)
 RP SEQUENCE OF 17-402 FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE-Placenta;
 RX MEDLINE-87105925; PubMed-3026837;
 RA Wu T.C., Kretzmer K.K.;
 RT "cDNA cloning and expression in E. coli of a plasminogen activator
 RT inhibitor (PAI) related to a PAI produced by Hep G2 hepatoma cell.";
 RL FEBS Lett. 210:11-16(1987).
 RN (12)
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
 RX MEDLINE-92114970; PubMed-1731226;

RA Mottonen J., Strand A., Syversky J., Sweet R.M., Danley D.E.,
 RA Geoghegan K.F., Gerard R.D., Goldsmith E.J.;
 RT "Structural basis of latency in plasminogen activator inhibitor-1.";
 RL Nature 355:270-273(1992).
 RN [13]
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
 RX MEDLINE-96003732; PubMed-7552714;
 RA Aertgeerts K., de Bondt H.L., de Kanter C.J., Declercq P.J.;
 RT "Mechanisms contributing to the conformational and functional
 RT flexibility of plasminogen activator inhibitor-1.";
 RL Nat. Struct. Biol. 2:891-897(1995).
 RN [14]
 RP X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS).
 RX MEDLINE-98298439; PubMed-9634700;
 RA Xue Y., Bjorkquist P., Inghardt T., Linschoten M., Musil D., Sjolin L.,
 RA Delum J.;
 RT "Interfering with the inhibitory mechanism of serpins: crystal
 RT structure of a complex formed between cleaved plasminogen activator
 RT inhibitor type 1 and a reactive-centre loop peptide.";
 RL Structure 6:627-636(1998).
 RN [15]
 RP X-RAY CRYSTALLOGRAPHY (2.99 ANGSTROMS).
 RX MEDLINE-99148119; PubMed-10368279;
 RA Sharp A.M., Stein P.E., Pannu N.S., Carrell R.W., Berkenpas M.B.,
 RA Glasburg D., Lawrence D.A., Read R.J.;
 RT "The active conformation of plasminogen activator inhibitor 1, a
 RT target for drugs to control fibrinolysis and cell adhesion.";
 RL Structure 7:111-118(1999).
 RN [16]
 RP VARIANT THR-15.
 RX MEDLINE-97337920; PubMed-9194591;
 RA Turkmen B., Schmitt M., Schmalteit B., Trommler P., Hell W.,
 RA Creutzburg S., Graeff H., Magdolen V.;
 RT "Mutational analysis of the genes encoding urokinase-type plasminogen
 RT activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer.";
 RL Electrophoresis 18:686-689(1997).
 CC [1]
 CC FUNCTION: THIS INHIBITOR ACTS AS "BAIT" FOR TISSUE PLASMINOGEN
 CC ACTIVATOR, UROKINASE, AND PROTEIN C. ITS RAPID INTERACTION WITH
 CC TPA MAY FUNCTION AS A MAJOR CONTROL POINT IN THE REGULATION OF
 CC FIBRINOLYSIS.
 CC [2]
 CC DISEASE: HIGH CONCENTRATIONS OF THIS PROTEIN HAVE BEEN ASSOCIATED
 CC WITH HUMAN THROMBOEMBOLIC DISEASE.
 CC [3]
 CC MISCELLANEOUS: PAI IS INACTIVATED BY PROTEOLYTIC ATTACK OF THE
 CC UROKINASE-TYPE (U-PA) AND THE TISSUE-TYPE (TPA), CLEAVING THE
 CC 369(R)-370(M) BOND.
 CC [4]
 CC MISCELLANEOUS: TWO TYPES OF PAI HAVE BEEN IDENTIFIED. PAI-1 IS AN
 CC ACID-STABLE GLYCOPROTEIN FOUND IN PLASMA AND PLATELETS AND IN
 CC ENDOTHELIAL, HEPATOMA, AND FIBROSARCOMA CELLS.
 CC [5]
 CC MISCELLANEOUS: VASCULAR ENDOTHELIAL CELLS MAY BE THE PRIMARY SITE
 CC OF SYNTHESIS OF PLASMA PAI.
 CC [6]
 CC SIMILARITY: BELONGS TO THE SERPIN FAMILY.
 CC [7]
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation.
 CC The European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isd-sib.ch/announce/>
 CC or send an email to license@isd-sib.ch).
 CC [8]
 CC EMBL: X04429; CAA28025.1; -
 CC EMBL: M14083; AAA60008.1; -
 CC EMBL: X04729; CAA28438.1; -
 CC EMBL: X04731; CAA28442.1; -
 CC EMBL: M16006; AAA60003.1; -
 CC EMBL: M22321; AAA60009.1; -
 CC EMBL: M22314; AAA60009.1; JOINED.
 CC EMBL: M22316; AAA60009.1; JOINED.
 CC EMBL: M22317; AAA60009.1; JOINED.
 CC EMBL: M22318; AAA60009.1; JOINED.
 CC EMBL: M22319; AAA60009.1; JOINED.
 CC EMBL: M22320; AAA60009.1; JOINED.

DR EMBL: X13323; -; NOT_ANNOTATED_CDS.
 DR EMBL: X13338; CAA11722.1; -
 DR EMBL: X13339; CAB51639.1; -
 DR EMBL: X13340; CAB51737.1; -
 DR EMBL: X13341; CAB51606.1; -
 DR EMBL: X13342; CAB51607.1; -
 DR EMBL: X13343; CAB51738.1; -
 DR EMBL: X13344; CAB51739.1; -
 DR EMBL: X13345; CAA11729.1; -
 DR EMBL: X13701; CAA11208.1; -
 DR EMBL: X03764; AAA60007.1; -
 DR EMBL: X04744; CAA28444.1; -
 DR EMBL: AC004876; AADA5828.1; -
 DR PIR: A26996; ITHUP1.
 DR PIR: S02551; S02551.
 DR PDB: 9PAI; 19-MAR-99.
 DR PDB: 1A7C; 23-MAR-99.
 DR PDB: 1B3K; 10-DEC-99.
 DR SWISS-2DPAGE; P05121; HUMAN.
 DR MIM; 173360; -
 DR InterPro: IPR000215; Serpin.
 DR Pfam; PF00079; serpin; 1.
 DR SMART; SM00093; SERPIN; 1.
 DR PROSITE; PS00284; SERPIN; 1.
 KW Serpin; serine protease inhibitor; Plasma; Plasminogen activation;
 KW Glycoprotein; Polymorphism; Signal; 3D-structure.
 FT SIGNAL 1 23
 FT CHAIN 1 402
 FT ACR SITE 24 370
 FT CARBOHYD 232 232
 FT CARBOHYD 288 288
 FT CARBOHYD 352 352
 FT VARIANT 15 15
 FT A->T.
 FT /FTID-VAR_007099.
 Query Match 100.0%; Score 29; DB 1; Length 402;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EITMD 6
 DB 373 EITMD 378
 RESULT 5
 ID PAI1_PIG STANDARD; PRT; 402 AA.
 AC P79335.
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE PLASMINOGEN ACTIVATOR INHIBITOR-1 PRECURSOR (PAI-1) (ENDOTHELIAL
 DE SERPINE1 OR PAI1 OR PLANH1.
 GN SERPINE1 OR PAI1 OR PLANH1.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-97206538; PubMed-9157595;
 RA Bjnens A.P., Knockaert I., Cousin E., Krutthof E.K.O., Declercq P.J.;
 RT "Expression and characterization of recombinant porcine plasminogen
 RT activator inhibitor-1.";
 RL Thromb. Haemost. 77:350-356(1997).
 RN [2]
 RP ERRATUM.
 RA Bjnens A.P., Knockaert I., Cousin E., Krutthof E.K.O., Declercq P.J.;
 RL Thromb. Haemost. 77:1046-1046(1997).
 CC [1]
 CC FUNCTION: THIS INHIBITOR ACTS AS "BAIT" FOR TISSUE PLASMINOGEN
 CC ACTIVATOR, UROKINASE, AND PROTEIN C. ITS RAPID INTERACTION WITH
 CC TPA MAY FUNCTION AS A MAJOR CONTROL POINT IN THE REGULATION OF
 CC FIBRINOLYSIS.

CC -1- MISCELLANEOUS: PAII IS INACTIVATED BY PROTEOLYTIC ATTACK OF THE
 CC UNOXYGENASE-TYPE (U-PA) AND THE TISSUE-TYPE (TPA), CLEAVING THE
 CC 369(R)-370(M) BOND (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Y11347; CAA72182.1; -
 CC HSSP: P01008; IATH.
 CC InterPro: IPR00215; Serpin.
 CC Pfam: PF00079; serpin. 1.
 CC SMART: SM00093; SERPIN. 1.
 CC PROSITE: PS00284; SERPIN. 1.
 CC Serpin, Serine protease inhibitor; Plasma; Plasminogen activation;
 CC Glycoprotein; Signal.
 CC SIGNAL 1 23 BY SIMILARITY.
 CC FT CHAIN 1 23 PLASMINOGEN ACTIVATOR INHIBITOR-1.
 CC FT ACT SITE 369 370 REACTIVE BOND.
 CC FT CARBOHYD 232 232 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 288 288 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC SEQUENCE 402 AA; 45450 MW; 13F0E5F48FE405 CRC64;

Query Match 100.0%; Score 29; DB 1; Length 402;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELMD 6
 DB 373 EELMD 378
 ID 111111
 ID CDNA2.YEAST STANDARD; PRT; 301 AA.
 AC 006702;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE CHITIN DEACETYLASE 2 PRECURSOR (EC 3.5.1.41).
 GN CDNA2 OR Y1R307W OR I2142.2.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=5288C / AB972;
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
 RA Favella A., Fulton L., Gattung S., Greco T., Kirsten J.,
 RA Kucaba T., Hallsworth K., Hawkins J., Hillier L., Jier M.,
 RA Johnson D., Johnston L., Langston Y., Larellie P., Le T.,
 RA Mardis E., Menezes S., Miller N., Nhan N., Pauley A., Peluso D.,
 RA Rifkin L., Riles L., Tatch A., Trevisan E., Vignati D.,
 RA Wilcox L., Wolfdmeyer F., Yandell M., Wilson R., Waterston R.;
 RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
 RN (12)
 RP CHARACTERIZATION.
 RX MEDLINE=97279228; PubMed=9133736;
 RA Mishra C., Semino C.E., McCreath K.J., de la Vega H., Jones B.J.,
 RA Specht C.A., Robbins P.W.;
 RT Cloning and expression of two chitin deacetylase genes of
 RT Saccharomyces cerevisiae.*;
 RL Yeast 13:327-336(1997).
 CC -1- FUNCTION: HYDROLYZES THE N-ACETAMIDO GROUPS OF N-ACETYL-D-
 CC GLUCOSAMINE RESIDUES IN CHITIN.

CC -1- CATALYTIC ACTIVITY: CHITIN + H(2)O -> CHITOSAN + ACETATE.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED EXCLUSIVELY DURING SPOROULATION.
 CC -1- SIMILARITY: TO OTHER POLYSACCHARIDE DEACETYLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U17247; AA67356.1; -
 CC SCD: S0004299; CDA2.
 CC InterPro: IPR002509; Polysac_deacet.
 CC Pfam: PF01522; Polysac_deacet. 1.
 CC Hydrolyase; Chitin degradation; Signal; Glycoprotein; Sporulation.
 CC SIGNAL 1 7 POTENTIAL.
 CC FT CHAIN 1 301 CHITIN DEACETYLASE 2.
 CC FT CARBOHYD 26 26 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 50 50 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 189 189 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC SEQUENCE 301 AA; 34642 MW; 72EE925B5E232 CRC64;

Query Match 93.1%; Score 27; DB 1; Length 301;
 Best Local Similarity 83.3%; Pred. No. 27;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELMD 6
 DB 243 EELMD 248
 ID 111111
 ID CYB5.YEAST STANDARD; PRT; 120 AA.
 AC P40312;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE CYTOCHROME B5.
 GN CYB5 OR YNL111C OR N1949.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=FL100;
 RX MEDLINE=94237477; PubMed=8181746;
 RA Trian G., Epinat J.-C., Rougeulle C., Cullin C., Pompon D.;
 RT Cloning and characterization of a yeast cytochrome b5-encoding gene
 RT which suppresses ketoconazole hypersensitivity in a MADH1-P-450
 RT reductase-deficient strain.*;
 RL Gene 142:123-127(1994).
 RN (12)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97245296; PubMed=9090055;
 RA de Antoni A., D'Angelo M., Dal Pero F., Sartorello F., Pandolfo D.,
 RA Pallavicini A., Lanfranchi G., Valle G.;
 RT The DNA sequence of cosmid 14-11b from chromosome XIV of
 RT Saccharomyces cerevisiae reveals an unusually high number of
 RT overlapping open reading frames.*;
 RL Yeast 13:261-266(1997).
 CC -1- FUNCTION: CYTOCHROME B5 IS A MEMBRANE BOUND HEMOPROTEIN WHICH
 CC OXYGENASES. IT PLAYS A ROLE IN FATTY-ACID DESATURATION AND IS ALSO
 CC INVOLVED IN SEVERAL STEPS OF THE SERUL BIOSYNTHESIS PATHWAY,
 CC PARTICULARLY IN THE 4-DEMETHYLATION OF THE 4,4'-DIMETHYL
 CC ZYMOSTROL.
 CC -1- SUBCELLULAR LOCATION: MICROSOMAL MEMBRANE. BOUND TO THE

CC CYTOPLASMIC SIDE OF THE ENDOPLASMIC RETICULUM (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B5 FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: L22494; AAA67468.1; -

DR EMBL: Z69382; CA93390.1; -

DR EMBL: Z71387; CA93390.1; -

DR PIR: S47919; S47919.

DR HSSP: P04166; 1B5M.

DR SCD: S0005055; CYB5.

DR InterPro: IPR001199; CYL_B5.

DR Pfam: PF00173; heme_1; 1.

DR PRINTS: PR00363; CYTOCHROME_B5.

DR PROSITE: PS00191; CYTOCHROME_B5_2; 1.

DR PROSITE: PS02255; CYTOCHROME_B5_2; 1.

DR Electron transport; Transmembrane; Heme; Microsome.

DM BINDING 37 37 HEME LIGAND (BY SIMILARITY).

FT BINDING 61 61 HEME LIGAND (BY SIMILARITY).

FT TRANSMEM 98 118 POTENTIAL.

FT CONFLICT 17 17 E -> Q (IN REF. 1).

SO SEQUENCE 120 AA; 13297 MW; 598EFA26730CAD19 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 120;

Best Local Similarity 83.3%; Pred. No. 17;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EELIIMD 6

DB 41 DEIIMD 46

RESULT 8

RBP6 ASPB7 STANDARD; PRT: 147 AA.

AC P42484; 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE DNA-DIRECTED RNA POLYMERASE, SUBUNIT 6 HOMOLOG (EC 2.7.7.6).

GN C147L.

OS African swine fever virus (strain BA71V) (ASFV).

OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae;

CC African swine fever-like viruses.

OX NCBI_TaxID=10498;

RN [1]

RP SEQUENCE FROM N.A.

RA Tanez R.J., Rodriguez J.M., Nogal M.L., Yuste L., Enriquez C.,

RA Rodriguez J.F., Vinuela E.;

RT "Analysis of the complete nucleotide sequence of African swine fever

RT virus.";

RL Virology 208:249-278(1995).

CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION

CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS

CC SUBSTRATES.

CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE - N PYROPHOSPHATE +

CC RNA(N).

CC -1- SIMILARITY: BELONGS TO THE ARCHAEABACTERIA RPOK / EUKARYOTIC RPB6

CC RNA POLYMERASE SUBUNIT FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: U18466; AAA65299.1; -

DR InterPro: IPR001725; RNA_pol_K_14ND.

DR Pfam: PF01192; RNA_pol_K; 1.

DR PROSITE: PS01111; RNA_POL_K_14KD; 1.

DR Transferrase; DNA-directed RNA polymerase; Transcription.

SO SEQUENCE 147 AA; 16691 MW; 6D0B6D77025CD45 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 147;

Best Local Similarity 83.3%; Pred. No. 21;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EELIIMD 6

DB 7 DEIIMD 12

RESULT 9

YE95_METJA STANDARD; PRT: 292 AA.

AC Q58890; 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE HYPOTHETICAL PROTEIN MJ1495.

DM MJ1495.

OS Methanococcus jannaschii.

OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;

CC Methanococcus

OX NCBI_TaxID=2190;

RN [1]

RP SEQUENCE FROM N.A.

RA STRAIN-JAL-1 / DSM 2661 / ATCC 43067;

RX MEDLINE=96337999; PubMed=8688087;

RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,

RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,

RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reick C.I.,

RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Nguyen D.,

RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,

RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,

RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,

RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;

RT "Complete genome sequence of the methanogenic archaeon, Methanococcus

RT jannaschii.";

RL Science 273:1058-1073(1996).

CC -1- SIMILARITY: STRONG, TO M. JANNASCHII MJ0137.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: U67590; AAB9506.1; -

DR TIGR: MJ1495; -

DR InterPro: IPR002825; DUF114.

DR Pfam: PF01972; DUF114; 1.

KM Hypothetical protein; Transmembrane; Complete proteome.

FT TRANSMEM 17 37 POTENTIAL.

SO SEQUENCE 292 AA; 33508 MW; BBA34C29B3A51246 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 292;

Best Local Similarity 83.3%; Pred. No. 45;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EELIIMD 6

DB 149 DEIIMD 154

RESULT 10
VG28_HAEIN STANDARD; PRT: 508 AA.
AC P44224;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE MO-LIKE PROPHAGE FLUO PROTEIN GP28.
GN H11500.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geophagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT *Whole-genome random sequencing and assembly of Haemophilus
RT Influenzae Rd *.
RL Science 269:496-512(1995).
CC -1- SIMILARITY: STRONG, TO PHAGE MU PROTEIN GP28.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U32826; AAC23150.1; -
DR TIGR: H11500; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 508 AA; 57190 MW; 5BFC6833E2P1242B CRC64;

Query Match 89.7%; Score 26; DB 1; Length 508;
Best Local Similarity 83.3%; Pred. No. 84;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EITIMD 6
DB 429 EITIMD 434

RESULT 11
YLN9_CAEEL STANDARD; PRT: 662 AA.
ID YLN9_CAEEL
AC Q09512;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHEICAL 76.7 KDA PROTEIN D2013.9 IN CHROMOSOME II.
GN D2013.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Rhabditidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Mortimore B., Matthews P.;

Submitted (DEC-1998) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 SET DOMAIN.
CC -1- SIMILARITY: STRONG, TO HUMAN KIAA0153.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: 247808; CAAB7778.1; -
DR EMBL: 247809; CAAB7778.1; JOINED.
DR EMBL: 247809; CAAB7783.1; -
DR EMBL: 247808; CAAB7783.1; JOINED.
DR Wormpep: D2013.9; CE01535.
DR InterPro: IPR001214; SET.
DR PROSITE: PS50280; SET: 1.
KW Hypothetical protein.
SQ SEQUENCE 662 AA; 76750 MW; 5044C76422145698 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 662;
Best Local Similarity 83.3%; Pred. No. 1;e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EITIMD 6
DB 249 EITIMD 254

RESULT 12
Y661_METJA STANDARD; PRT: 762 AA.
ID Y661_METJA
AC Q58371;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHEICAL MCM-TYPE PROTEIN MJ0961.
GN MJ0961.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Belch C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geophagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts R.M., Hurt M.A., Kaine B.P., Borodovsky M.,
RA Klein H.-P., Fraser C.M., Smith H.O., Weese C.R., Venter J.C.;
RT *Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii *.
RL Science 273:1058-1073(1996).
CC -1- SIMILARITY: BELONGS TO THE MCM FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U67539; AAB98963.1; -
DR TIGR: MJ0961; -
DR InterPro: IPR001208; MCM.

DR Pfam: PF00493; MCM, 1.
 DR SMART: SM0050; MCM, 1.
 DR PROSITE: PS00847; MCM, 1.
 DR PROSITE: PS50051; MCM, 2, 1.
 KW Hypothetical protein; Transcription regulation; DNA-binding;
 KW DNA replication; Cell cycle; ATP-binding; Complete proteome.
 FT DOMAIN 334 542 MCM.
 FT NP_BIND 384 391 ATP (POTENTIAL).
 SO SEQUENCE 762 AA; 87526 MW; 3323509057464C12 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 762;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIIIND 6
 |||||
 Db 64 EIIIND 69

RESULT 13
 YHGF_ECOLI STANDARD; PRT; 773 AA.
 ID YHGF_ECOLI
 AC P46837; P76689;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE PROTEIN YHGF.
 GN YHGF OR B3407.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE-97426617; PubMed-9278503;
 RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-VIDES J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12."
 RL Science 277:1453-1474(1997).
 RN [2]
 RP IDENTIFICATION BY MASS SPECTROMETRY.
 RX MEDLINE-99420866; PubMed-10493123;
 RA Fountoulakis M., Takacs M.-F., Berndt P., Langen H., Takacs B.;
 RT "Enrichment of low abundance proteins of Escherichia coli by
 RT hydroxyapatite chromatography."
 RL Electrophoresis 20:2181-2195(1999).
 CC -1- SIMILARITY: STRONG TO H. INFLUENZAE H10568.
 CC -1- SIMILARITY: CONTAINS 1 'SI MOTIF' DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: U18997; AAA58204.1; ALT_FRAME.
 DR EMBL: U18997; AAA58205.1; ALT_FRAME.
 DR EMBL: AE000416; AAC76437.1; ALT_INIT.
 DR HSRF: P05055; ISRO.
 DR Ecocore: EGI2932; YHGF.
 DR InterPro: IPR003029; SI.
 DR Pfam: PF00575; SI; 1.
 DR SMART: SM00316; SI; 1.
 KW RNA-binding; Complete proteome.
 FT DOMAIN 651 720 SI MOTIF.
 FT CONFLICT 754 755 OP -> HA (IN REF. 1; AAA58205).
 SO SEQUENCE 773 AA; 85119 MW; EA54D9ED952A8229 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 773;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIIIND 6
 |||||
 Db 255 EIIIND 260

RESULT 14
 CDRB_CLOAB STANDARD; PRT; 139 AA.
 ID CDRB_CLOAB
 AC Q45811;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE CHEMICAL-DAMAGING AGENT RESISTANCE PROTEIN B.
 OS Clostridium acetobutylicum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ABKN8;
 RX MEDLINE-95038273; PubMed-7765497;
 RA Azeedong H., Reysset G.;
 RT "Cloning and sequencing of a chromosomal fragment from Clostridium
 RT acetobutylicum strain ABKN8 conferring chemical-damaging agents and
 RT UV resistance to B. coli-like strains."
 RL Curr. Microbiol. 29:229-235(1994).
 CC -1- FUNCTION: NOT KNOWN; COULD CONFER METHYL METHANE SULFONATE (MMS),
 CC MITOMYCIN C (MC), AND UV RESISTANCE.
 CC -1- SIMILARITY: BELONGS TO THE CAPAB / TENDENZ FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X74918; CNA52879.1; -.
 DR InterPro: IPR003325; Teryd.
 DR Pfam: PF02342; Teryd; 1.
 SO SEQUENCE 139 AA; 15429 MW; 51D5656D0CAB9827 CRC64;

Query Match 86.2%; Score 25; DB 1; Length 139;
 Best Local Similarity 83.3%; Pred. No. 35;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIIIND 6
 |||||
 Db 37 EIIIND 42

RESULT 15
 PPB6_ASFM2 STANDARD; PRT; 139 AA.
 ID PPB6_ASFM2
 AC Q08399;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE DNA-DIRECTED RNA POLYMERASE, SUBUNIT 6 HOMOLOG (EC 2.7.7.6).
 GN PMW-CL.
 OS African swine fever virus (isolate Malawi L1 20/1) (ASFV).
 OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae;
 OC African swine fever-like viruses.
 OX NCBI_TaxID=10500;
 RN [1]

RP SEQUENCE FROM N.A. PubMed-8332503;
 RX MEDLINE-93324382; Rock D.L.;
 RA Lu Z., Kutish G.F., Susman M.D.,
 RT "An African swine fever virus gene with a similarity to eukaryotic
 RNA polymerase subunit 6."
 RL Nucleic Acids Res. 21:2940-2940(1993).
 CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
 OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
 SUBSTRATES.
 CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE - N PYROPHOSPHATE +
 RNA(N).
 CC -1- SIMILARITY: BELONGS TO THE ARCHAEABACTERIA RPOK / EUKARYOTIC RPB6
 CC RNA POLYMERASE SUBUNIT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L15411; AAA42731.1;
 DR InterPro: IPR001725; RNA_POL_K_14KD.
 DR Pfam: PF01192; RNA_POL_K_1.
 DR PROSITE: PS01111; RNA_POL_K_14KD; 1.
 KW Transferase; DNA-directed RNA polymerase; Transcription.
 SQ SEQUENCE 139 AA; 15713 MW; 0E96156D214E01D CRC64;

Query Match 86.2%; Score 25; DB 1; Length 139;
 Best Local Similarity 66.7%; Pred. No. 35;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EEDMD 6
 I:::|||
 DB 7 EDVIMD 12

Search completed: March 18, 2002, 09:52:26
 Job time: 295 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compen Ltd.

OW protein - protein search, using sw model

Run on: March 18, 2002, 09:51:56 ; Search time 78.82 Seconds

(without alignments)
11.135 Million cell updates/sec

Title: US-09-544-665-3
Perfect score: 29
Sequence: 1 EELMD 6

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 473505 seqs, 14627329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	100.0	284	10 082397	082397 arabidopsis
2	29	100.0	525	2 051223	051223 borrelia bu
3	28	96.6	460	1 09HNB1	09HNB1 methanococ
4	27	93.1	120	2 09A256	09A256 caulobacter
5	27	93.1	319	10 09FLU5	09FLU5 arabidopsis
6	27	93.1	574	5 09XWJ3	09XWJ3 caenorhabdi
7	27	93.1	1052	10 09LW13	09LW13 arabidopsis
8	26	88.7	152	11 09COW6	09COW6 mus musculu
9	26	88.7	163	4 09T210	09T210 homo sapien
10	26	88.7	163	11 09D385	09D385 mus musculu
11	26	88.7	175	11 09D970	09D970 mus musculu
12	26	88.7	180	5 09BMB5	09BMB5 ancylostoma
13	26	88.7	204	4 09N283	09N283 homo sapien
14	26	88.7	232	2 09F0U6	09F0U6 desulfiloba
15	26	88.7	274	1 059224	059224 pyrococcus
16	26	88.7	274	1 09V118	09V118 pyrococcus
17	26	88.7	284	2 09X018	09X018 thermotoga
18	26	88.7	288	2 066996	066996 aquifex aeo
19	26	88.7	335	2 09AN58	09AN58 bradyrhizob

20	26	89.7	346	10 09SHB4	09SHB4 arabidopsis
21	26	89.7	424	10 041109	041109 phaseolus v
22	26	89.7	446	4 09BQ67	09BQ67 homo sapien
23	26	89.7	695	5 09V5N5	09V5N5 drosophila
24	26	89.7	739	10 09M9G4	09M9G4 arabidopsis
25	26	89.7	774	5 09VJ73	09VJ73 drosophila
26	26	89.7	912	5 09NGP1	09NGP1 dictyosteli
27	26	89.7	3885	5 09N5J3	09N5J3 caenorhabdi
28	26	86.2	107	2 09P1D2	09P1D2 unclassified
29	25	86.2	120	2 09MXP5	09MXP5 thermotoga
30	25	86.2	128	12 072563	072563 human immun
31	25	86.2	142	2 09F354	09F354 streptomyce
32	25	86.2	222	1 09UX10	09UX10 sulfolobus
33	25	86.2	237	2 09CFE23	09CFE23 lactococcus
34	25	86.2	243	7 031362	031362 brachydanio
35	25	86.2	245	2 09WYMS	09WYMS thermotoga
36	25	86.2	277	2 09EXX1	09EXX1 enterococcu
37	25	86.2	308	10 09XFR7	09XFR7 arabidopsis
38	25	86.2	308	10 09RKP7	09RKP7 arabidopsis
39	25	86.2	357	10 09ZNX0	09ZNX0 nicotiana t
40	25	86.2	361	1 09U250	09U250 pyrococcus
41	25	86.2	365	1 09HHR2	09HHR2 halobacteri
42	25	86.2	389	1 027747	027747 methanobact
43	25	86.2	411	2 09L6G2	09L6G2 lactococcus
44	25	86.2	419	5 044714	044714 caenorhabdi
45	25	86.2	429	10 09FGC3	09FGC3 arabidopsis

ALIGNMENTS

RESULT 1
ID 082397 PRELIMINARY: PRT; 284 AA.
AC 082397;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE HYPOTHETICAL 32.6 KDA PROTEIN.
GN T30120.2
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucotyledons II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,
RA Shen M., Ronning C.M., Fraser C.M., Somerville C.R., Venter J.C.;
RT "Arabidopsis thaliana chromosome II BAC T30120 genomic sequence.";
DR Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC005617; AAC63587.1;
DR InterPro: IPR001092; HLH_dim.
DR Pfam: PF00010; HLH; 1.
DR SMART: SM00353; HLH; 1.
DR SMART: SM00353; HLH; 1.
SQ SEQUENCE 284 AA: 32624 MW: 239AD2B6D7134AA4 CRC64;

Query Match 100.0%; Score 29; DB 10; Length 284;

Best Local Similarity 100.0%; Pred. No. 73;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EELMD 6

DB 90 EELMD 95

RESULT 2

ID 051223 PRELIMINARY: PRT; 525 AA.

AC 051223;
DT 01-JUN-1998 (TREMBLrel. 06, Created)

DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
 DT 01-NOV-1998 (TREMblrel. 08, Last annotation update)
 DE HYPOTHETICAL 61.7 KDA PROTEIN.
 GN BB0205.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 CX NCBI_TaxId=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 35210 / B31;
 RX MEDLINE-98065943; PubMed-9403685;
 RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
 RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
 RA Dougherty B., Tomb J.F., Fleischmann R.D., Richardson D.,
 RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
 RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
 RA Uterback T., Matthey L., McDonald L., Artlisch P., Bowman C.,
 RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
 RA Smith H.O., Venter J.C.,
 RT Genome sequence of a Lyme disease spirochete, Borrelia
 burgdorferi.
 RL Nature 390:580-586(1997).
 DR EMBL; AE001131; AAC66601.1; .
 DR TIGR; BB0205; .
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 525 AA: 61743 MW: 837F0B2A35EB25D CRC64;

Query Match 100.0%; Score 29; DB 2; Length 525;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIIMD 6
 DB 461 EELIIMD 466

RESULT 3
 ID 09HHB1 PRELIMINARY; PRT: 460 AA.
 AC 09HHB1;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE PROLYL-TRNA SYNTHETASE (EC 6.1.1.15).
 OS Methanococcus maripaludis.
 OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
 OC Methanococcus.
 CX NCBI_TaxId=39152;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Stathopoulos C., Li T., Longman R., Voithnecht U.C., Becker H.D.,
 RA Ibbas M., Soll D.;
 RT "A single class II synthetase specifies two amino acids in archaeal
 translation."
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF197899; AAC28517.1; .
 DR InterPro: IPR002316; TRNA-synt_pro.
 DR PRINTS; PRO1046; TRNASYNTHPRO.
 KW Aminoacyl-tRNA synthetase; Ligase.
 SQ SEQUENCE 460 AA: 53292 MW: DA8364435B808E44 CRC64;

Query Match 96.6%; Score 28; DB 1; Length 460;
 Best Local Similarity 83.3%; Pred. No. 2.1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIIMD 6
 DB 292 EELIIMD 297
 RESULT 4

O9A256
 ID 09A256 PRELIMINARY; PRT: 120 AA.
 AC 09A256;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DE HYPOTHETICAL PROPEIN CC3710.
 GN CC3710.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
 OC Caulobacter.
 CX NCBI_TaxId=69394;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21173698; PubMed-11259647;
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Eelsen J., Helgelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Potocka I., Nelson M.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 RA Debey R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
 RA Uterback T., Tran K., Wolf A., Yamathayan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of Caulobacter crescentus."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 DR EMBL; AE006028; AAK25672.1; .
 DR TIGR; CC3710; .
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 120 AA: 12483 MW: B30E9BCA06F9283A CRC64;

Query Match 93.1%; Score 27; DB 2; Length 120;
 Best Local Similarity 66.7%; Pred. No. 86;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIIMD 6
 DB 13 EELIIMD 18

RESULT 5
 ID 09FLU5 PRELIMINARY; PRT: 319 AA.
 AC 09FLU5;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE GENOMIC DNA, CHROMOSOME 5, TAC CLONE:K18P6.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
 CX NCBI_TaxId=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA;
 RX MEDLINE-98290546; PubMed-9628582;
 RA Sato S., Kaneo F., Kotani H., Nakamura Y., Asamizu E., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. IV.
 RT Sequence features of the regions of 1,456,315 bp covered by nineteen
 RT physically assigned P1 and TAC clones."
 RL DNA Res. 5:41-54(1998).
 DR EMBL; AB010068; BAB11210.1; .
 DR InterPro: IPR001934; HMA.
 DR InterPro: IPR001230; Prenyltn.
 DR Pfam: PF00403; HMA; 2
 DR PROSITE; PS00294; PRENYLATION; UNKNOWN; 1
 SQ SEQUENCE 319 AA: 35350 MW: 5A97CA58833BCDE0 CRC64;

Query Match 93.1%; Score 27; DB 10; Length 319;
 Best Local Similarity 66.7%; Pred. No. 2.4e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMD 6
11:11
Db 84 EELVMD 89

RESULT 6

O9XWJ3 PRELIMINARY: PRT: 574 AA.

AC O9XWJ3;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Y5H1A.6 PROTEIN.
GN Y5H1A.6
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Smye R.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Aliscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
RA Gardner M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightling J., Lloyd C., McMurtry A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Woldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans";
RL Nature 368:32-38(1994).
DR EMBL: AL032644; CAA21671.1; -
DR InterPro: IPR000822; Znf_C2H2.
DR Pfam: PF00096; Znf_C2H2; 1.
DR SMART: SM00355; Znf_C2H2; 1.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
DR PROSITE: PS00157; ZINC_FINGER_C2H2_2; 1.
KW DNA-binding; zinc-finger.
SQ SEQUENCE 574 AA; 65776 MW; 91466762ADDAE5F0 CRC64;

Query Match 93.1%; Score 27; DB 5; Length 574;
Best Local Similarity 66.7%; Pred. No. 4.6e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMD 6
11:11
Db 130 EELVMD 135

O9LJL3 PRELIMINARY: PRT: 1052 AA.

AC O9LJL3;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE ZINC METALLOPROTEINASE (INSULINASE FAMILY).
OS Arabidopsis thaliana (mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN-COLUMBIA;
RA Kaneo T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RX PubMed=10907853;
RA Nakamura Y.;

RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the regions of 4,251,695 bp covered by niney P1,
RT TAC and BAC clones";
RL DNA Res. 7:217-221(2000).
DR EMBL: AP00419; BAB02957.1; -
KW Protease; Metalloprotease.
SQ SEQUENCE 1052 AA; 117654 MW; 40FB31F50BFE8A9B CRC64;

Query Match 93.1%; Score 27; DB 10; Length 1052;
Best Local Similarity 83.3%; Pred. No. 8.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMD 6
11:11
Db 434 EELVMD 439

RESULT 8

O9COW6 PRELIMINARY: PRT: 152 AA.

AC O9COW6;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE 1700027H16RIK PROTEIN.
GN 1700027H16RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TESTIS.
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinaigawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamahata I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Kono H., Balderelli R., Barsh G.,
RA Blake J., Botfield D., Boujunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gaitanaris M.P.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 403:685-690(2001).
DR EMBL: AK00646; BAB24686.1; -
DR EMBL: AK006418; BAB24578.1; -
DR MGI: MGI:191333; 1700027H16RIK.
SQ SEQUENCE 152 AA; 17571 MW; 4DE5F01744ED1E58 CRC64;

Query Match 89.7%; Score 26; DB 11; Length 152;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EELIMD 6
DB 18 EDIMD 23

RESULT 9

Q9Y2Y0 PRELIMINARY; PRT: 163 AA.

AC 09Y2Y0; 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RT "The Arf-like 2 (Arf2) binding protein BART1: purification, cloning,
and initial characterization."
RL Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.
RC [2]
RP SEQUENCE FROM N.A.
RC TISSUE=NEUROBLASTOMA;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF126062; AAD20633.1; -
DR EMBL; BC003087; AA03087.1; -
SQ SEQUENCE 163 AA; 18822 MW; E35B5AC73FC1FEC CRC64;

Query Match 89.7%; Score 26; DB 4; Length 163;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EELIMD 6
DB 29 EDIMD 34

Q9D385 PRELIMINARY; PRT: 163 AA.

AC 09D385; 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE 6310544B05R1K PROTEIN (SIMILAR TO BINDER OF ARL TWO).
GN 6310544B05R1K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=MEDULLA OBLONGATA;
RA MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,
Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
Kadota K., Matsuda H. A., Ashburner M., Batalov S., Casavant T.,
Raisz L. M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
Schriml L. M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M. F.,
Brownstein M. J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
Gustlich S., Hill D., Hofmann M., Hume D. A., Kamiya M., Lee N. H.,
Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
Suzuki H., Toyo-Oka K., Wang K. H., Weitz C., Whittaker C., Wilming L.,
Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai J., Kohsaki S.,
Hayashizaki Y.;
Functional annotation of a full-length mouse cDNA collection.
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC Strausberg R.;
RL Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AK018329; BAB31126.1; -
DR EMBL; BC002331; AA02121.1; -
DR MGI; MGI:1923421; 6310544B05R1K.
SQ SEQUENCE 163 AA; 18754 MW; AE489178C038225 CRC64;

Query Match 89.7%; Score 26; DB 11; Length 163;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EELIMD 6
DB 29 EDIMD 34

Q9D9T0 PRELIMINARY; PRT: 175 AA.

AC 09D9T0; 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE 1700029M23R1K PROTEIN.
GN 1700029M23R1K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TESTIS;
RA MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,
Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
Kadota K., Matsuda H. A., Ashburner M., Batalov S., Casavant T.,
Raisz L. M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
Schriml L. M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M. F.,
Brownstein M. J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
Gustlich S., Hill D., Hofmann M., Hume D. A., Kamiya M., Lee N. H.,
Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
Suzuki H., Toyo-Oka K., Wang K. H., Weitz C., Whittaker C., Wilming L.,
Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai J., Kohsaki S.,
Hayashizaki Y.;
Functional annotation of a full-length mouse cDNA collection.
RL Nature 409:685-690(2001).
DR EMBL; AK006511; BAB24626.1; -
DR MGI; MGI:1916746; 1700029M23R1K.
SQ SEQUENCE 175 AA; 20552 MW; 15A8B6A38F87706E CRC64;

Query Match 89.7%; Score 26; DB 11; Length 175;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EELIMD 6
DB 11 EDIMD 6

Db 133 EELVMD 138

RESULT 12

Q9BMB5 PRELIMINARY; PRT; 180 AA.
 AC Q9BMB5; 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE CATHEPSIN B-LIKE PRECURSOR PROTEIN (FRAGMENT).
 OS Ancylostoma ceylanicum.
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida; Ancylostomatidae; Ancylostomatidae; Ancylostomatidae; Ancylostoma.
 NCBI_TaxID=53326;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Harrison L.M., Cappello M.;
 RT Isolation of a cathepsin B-like sequence from adult Ancylostoma ceylanicum.
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF339098; AAK09441.1;
 KW Signal.
 FT SIGNAL 1 13 POTENTIAL.
 FT NON_TER 180
 SQ SEQUENCE 180 AA: 20099 MW; F3729B7CC683601 CRC64;

Query Match 89.7%; Score 26; DB 5; Length 180;
 Best Local Similarity 66.7%; Pred. No. 2.3e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EELVMD 6
 Db 78 EELVMD 83

RESULT 13

Q9N2B3 PRELIMINARY; PRT; 204 AA.
 AC Q9N2B3; 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE UNCHARACTERIZED BONE MARROW PROTEIN BM039.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE-BONE MARROW.
 RA Zhao M., Gu J., Li N., Peng Y., Han Z., Chen Z.;
 RT A novel gene expressed in human bone marrow.
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF217515; AAF67626.1;
 SQ SEQUENCE 204 AA: 24053 MW; 39E4B5F78E020260 CRC64;

Query Match 89.7%; Score 26; DB 4; Length 204;
 Best Local Similarity 83.3%; Pred. No. 2.6e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EELVMD 6
 Db 190 EELVMD 195

RESULT 14

Q9F0U6 PRELIMINARY; PRT; 232 AA.
 AC Q9F0U6; 01-MAR-2001 (TrEMBLrel. 16, Created)

DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE HYPOTHETICAL 26.9 KDA PROTEIN.
 OS Desulfotobacterium chlororespirans.
 CC Bacteria; Firmicutes; Bacillus/Clostridium group; Hellobacterium group; Desulfotobacterium.
 NCBI_TaxID=51616;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-CO23;
 RA Davis J.K., Tiedje J.M.;
 RT Sequence of o-chlorophenol reductive dehalogenase from Desulfotobacterium chlororespirans Co23.
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF204275; AAG3484.1;
 DR InterPro: IPR000595; CNMP_binding.
 DR InterPro: IPR001808; HTH_CRP.
 DR SMART: SM00100; CNMP.
 DR SMART: SM00419; HTH_CRP.
 DR PROSITE: PSS0042; CNMP_BINDING_3; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 232 AA: 26915 MW; DC268444CA054712 CRC64;

Query Match 89.7%; Score 26; DB 2; Length 232;
 Best Local Similarity 83.3%; Pred. No. 3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EELVMD 6
 Db 120 EELVMD 125

RESULT 15

Q59224 PRELIMINARY; PRT; 274 AA.
 AC Q59224; 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE 274AA LONG HYPOTHETICAL AUTOANTIGEN LIKE PROTEIN.
 CN PH1548.
 OS Pyrococcus horikoshii.
 CC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
 NCBI_TaxID=53953;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-OT3;
 RX MEDLINE-98344137; PubMed-9679194;
 RA Kawarabayashi Y., Sawada M., Horikawa H., Hata Y., Hino Y., Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y., Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y., Kashi K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K., Masuchi Y., Shizuya H., Kikuchi H.;
 RT Complete sequence and gene organization of the genome of a hyperthermophilic archaeobacterium, Pyrococcus horikoshii OT3.
 RL DNA Res. 5:55-76(1998).
 DR EMBL; AP000006; BAA30660.1;
 DR InterPro: IPR001247; 3_EXONASE.
 DR Pfam: PF01138; RNase_PH; 1.
 KW Complete proteome.
 SQ SEQUENCE 274 AA: 29994 MW; 29168C0C76B0490 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 274;
 Best Local Similarity 66.7%; Pred. No. 3.6e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EELVMD 6
 Db 213 EELVMD 218

Tue Mar 19 11:00:05 2002

us-09-544-665-3.rpt

Page 6

Search completed: March 18, 2002, 09:51:57
Job time: 286 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:48:56 ; Search time 81.32 seconds

(without alignments)
5.465 Million cell updates/sec

Title: US-09-544-665-4

Perfect score: 29

Sequence: 1 REIMD 6

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_1101.*
1: /SID8/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID8/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID8/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID8/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID8/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID8/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID8/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID8/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID8/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID8/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID8/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID8/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID8/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID8/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID8/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID8/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID8/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID8/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID8/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID8/gcgdata/geneseq/geneseq/AA2000.DAT.*
21: /SID8/gcgdata/geneseq/geneseq/AA2001.DAT.*
22: /SID8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	100.0	6	20	AAV01813
2	29	100.0	379	11	AAV07245
3	29	100.0	402	13	AAV23812
4	27	93.1	340	19	AAV98747
5	27	93.1	758	22	AAU03391
6	27	93.1	1245	22	AAV85015
7	26	89.7	259	22	AAV72392
8	25	86.2	58	17	AAV83032
9	25	86.2	58	20	AAV33889
10	25	86.2	102	21	AAV5542
11	25	86.2	102	21	AAV5543

12	25	86.2	174	13	AAV21412	NADH dehydrogenase
13	25	86.2	212	19	AAV6073	S. pneumoniae dext
14	25	86.2	212	21	AAV08637	Amino acid sequenc
15	25	86.2	212	21	AAV61763	Streptococcus pneu
16	25	86.2	351	21	AAV84439	Amino acid sequenc
17	25	86.2	351	22	AAV83497	Human protein sequ
18	25	86.2	351	22	AAV63050	Human fibronucleot
19	25	86.2	521	22	AAV32728	Corynebacterium prote
20	25	86.2	521	22	AAV93355	Corynebacterium 91
21	25	86.2	1522	20	AAV93302	Human BAI3 protein
22	25	86.2	1522	21	AAV23601	Human secreted pro
23	25	86.2	6	20	AAV01812	Peptide used to re
24	24	82.8	33	22	AAV42738	Human alpha-1-anti
25	24	82.8	33	22	AAV92200	Thrombin inhibitor
26	24	82.8	35	21	AAV69965	Core polypeptide f
27	24	82.8	35	21	AAV69966	Core polypeptide f
28	24	82.8	35	21	AAV69967	Core polypeptide f
29	24	82.8	35	21	AAV69968	Core polypeptide f
30	24	82.8	35	21	AAV69969	Core polypeptide f
31	24	82.8	35	21	AAV69970	Core polypeptide f
32	24	82.8	35	21	AAV90001	Core polypeptide f
33	24	82.8	35	22	AAV8386	Core polypeptide T
34	24	82.8	35	22	AAV8387	Core polypeptide T
35	24	82.8	35	22	AAV8388	Core polypeptide T
36	24	82.8	35	22	AAV8394	Core polypeptide T
37	24	82.8	35	22	AAV8395	Core polypeptide T
38	24	82.8	35	22	AAV8399	Core polypeptide T
39	24	82.8	35	22	AAV8402	Core polypeptide T
40	24	82.8	38	21	AAV69855	FIV gp36 envelope
41	24	82.8	38	21	AAV69856	FIV gp36 envelope
42	24	82.8	38	21	AAV69857	FIV gp36 envelope
43	24	82.8	38	21	AAV69858	FIV gp36 envelope
44	24	82.8	38	21	AAV69859	FIV gp36 envelope
45	24	82.8	38	21	AAV69860	FIV gp36 envelope

ALIGNMENTS

RESULT 1	
ID	AAV01813 standard; peptide: 6 AA.
XX	AAV01813;
AC	AAV01813;
XX	29-JUN-1999 (first entry)
DT	
XX	Peptide used to regulate scupa.
DE	
XX	Abnormal cell migration; PAI-1-dependent cell adhesion; scupa clearance;
KW	pathological cell migration; angiogenesis; organogenesis; ovulation;
KW	Inflammation; cancer; tumor cell invasion; metastasis; atherosclerosis.
KW	
XX	Synthetic.
OS	
XX	WO920295-A1.
PN	
XX	29-APR-1999.
PD	
XX	15-OCT-1998; 98WO-0521800.
PP	
XX	17-OCT-1997; 97US-0062274.
PR	
XX	(UYPE-) UNTV. PENNSYLVANIA.
XX	
PA	Clines D, Higaiz AA;
XX	
PI	WPI; 1999-288168/24.
XX	
DR	Peptide composition
XX	
PT	
XX	Example 1; Page 31; 63pp; English.
PS	
XX	

CC Peptides AAR01812-14 exemplify the peptide of the invention, which is
 CC used in a peptide composition to affect a biological process
 CC characterized by abnormal cell migration through a physiological barrier,
 CC to inhibit PAI-1-dependent adhesion of a cell to a tissue, to promote
 CC clearance of scuba from the surface of a mammalian cell, and to impede
 CC pathological migration of a cell in a mammal. The composition can be used
 CC to affect a biological process characterized by abnormal cell migration
 CC through a physiological barrier, where the process is selected from
 CC angiogenesis, organogenesis, ovulation, inflammation, cancer, tumor cell
 CC invasion and metastasis, and atherosclerosis.

XX
 SQ Sequence 6 AA;

Query Match 100.0%; Score 29; DB 20; Length 6;
 Best Local Similarity 100.0%; Pred. No. 4.3e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 REIMD 6
 |||||
 Db 1 reimd 6

RESULT 2
 AAR07245
 ID AAR07245 standard; protein: 379 AA.

XX AAR07245;

DT 24-JAN-1991 (first entry)

DE PAI-1(E350->R).

XX Tissue plasminogen activator; light chain; mutant; serine protease;
 KM chymotrypsin; salt bridge; serpin; inhibitor; t-PA.

XX Homo sapiens.

XX MO9010649-A.

XX 20-SEP-1990.

XX 01-MAR-1990; 90MO-US00947.

XX 13-NOV-1989; 88US-0434748.

XX 06-MAR-1989; 89US-0319212.

XX (TEXA) UNIV TEXAS SYSTEM.

XX Sambrook JF, Madison EL, Goldsmith EJ, Gething MH;

PI Gerard RD;

DR WPI: 1990-304988/40.

XX Serine protease mutant of the chymotrypsin superfamily -
 PT compsn. including it and genes encoding it, resistant to inhibition
 PT by PAI-1.

PS Claim 34; Page 60; 75pp; English.

XX The sequence represents AAs 1-380 of the PAI-1 protein chain.
 CC The mutant was obtained by site directed mutagenesis of the wild type
 CC gene. It has an Arg at posn. 350 replacing the Glu residue which
 CC is thought to form a contact with the t-PA Arg residue at posn.
 CC 304. The binding is modelled on that of the complex formed between
 CC trypsin and BPTI where the Tyr39 of trypsin forms a Van der Waal's
 CC contact with Ile19 of BPTI. In the mutant the Glu residue has been
 CC replaced by and Arg residue, restoring the electrostatic
 CC interaction that was disrupted by the construction of the Arg304->
 CC Glu mutation in t-PA (see AAR07240). This serpin mutant is therefore
 CC complementary to the mutant t-PA. The mutant serpin is a more
 CC active inhibitor than its parent protein.
 CC See also AAR07239-46.

XX
 SQ Sequence 379 AA;

Query Match 100.0%; Score 29; DB 11; Length 379;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 REIMD 6
 |||||
 Db 350 reimd 355

RESULT 3
 AAR23812
 ID AAR23812 standard; protein: 402 AA.

XX AAR23812;

DT 30-OCT-1992 (first entry)

DE PAI-1 (Arg 350) mutant.

XX Tissue plasminogen activator; mutant; serine protease;
 KM zymogen-like; PAI-1; plasminogen activator inhibitor; serpin;
 KM coagulation; chymotrypsin superfamily.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1-23

FT Protein /label= leader

FT /label= mature_PAI-1

XX MO9206203-A.

XX 16-APR-1992.

XX 10-SEP-1991; 91MO-US06366.

XX 28-SEP-1990; 90US-0589554.

XX (TEXA) UNIV TEXAS SYSTEM.

XX Gerard RD, Gething MH, Goldsmith EJ, Madison EL, Sambrook JF;

DR WPI: 1992-150888/18.

XX New zymogen plasminogen activator mutants - resistant to
 PT inhibition by cognate serpin inhibitors in plasma and platelets,
 PT for myocardial infarction

PS Example 2; Page 55; 82pp; English.

XX This PAI-1 mutant has Glu 350 of wild-type mature PAI-1
 CC substituted by Arg. This substitution was chosen to selectively
 CC alter the negatively-charged Glu residue to a positively-charged
 CC Arg residue and promote potential interactions with the negatively
 CC charged Glu 304 residue in the PAI-1-resistant tPA mutant
 CC tPA(Glu304) (see AAR23804). The mutant PAI-1 showed an increased rate
 CC constant of interaction with tPA(Glu304) c.f. wild-type PAI-1.
 CC The sequence given as AAR23813 does not appear in the specification
 CC but has been compiled from the known PAI-1 sequence (see MO9013648)
 CC and the description given in the specification.
 CC See also AAR23801-R23813.

XX Sequence 402 AA;

Query Match 100.0%; Score 29; DB 13; Length 402;
 Best Local Similarity 100.0%; Pred. No. 43;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RETIND 6
 |||||
 DB 373 retind 378

RESULT 4

AAW98747
 ID AAW98747 standard; Protein; 340 AA.

AC AAW98747;

DT 31-MAR-1999 (first entry)

DE H. pylori GHP0 1033 protein.

KN GHP0 protein; Helicobacter infection; gastroduodenal disease; gastritis;
 KM peptic ulcer disease.

OS Helicobacter pylori.

PN MO9843478-A1.

PD 08-OCT-1998.

PF 01-APR-1998; 98MO-US06371.

PR 29-JUL-1997; 97US-0902615.

PR 01-APR-1997; 97US-083457.

PR 24-JUN-1997; 97US-0881227.

PA (HUMA-) HUMAN GENOME SCI INC.

PA (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.

PI A1-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;

DR WPI: 1998-542293/46.

DR N-PSDB; AAX14466.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

DE Pyrenophora teres partial CPS1 peptide synthetase.
 XX
 KM CPS1; peptide synthetase; peptide toxin; fungal pathogen;
 KM corn crop infection.

OS Pyrenophora teres.

FT Key Location/Qualifiers

FT Region 73..86 /note= "Peptide synthetase motif, core 3"

FT Region 118..125 /note= "Peptide synthetase motif, core 4"

FT Region 135..151 /note= "Peptide synthetase motif, core 5"

FT Region 217..225 /note= "Cyclization domain motif"

FT Region 221..224 /note= "Peptide synthetase motif, core 6"

FT Region 424..435 /note= "Peptide synthetase motif, core 1"

FT Region 524..543 /note= "Peptide synthetase motif, core 2"

FT Region 755..761 /note= "Peptide synthetase motif, core 3"

XX MO200138489-A2.

XX 31-MAY-2001.

XX 22-NOV-2000; 2000MO-US32227.

XX 23-NOV-1999; 99US-0448215.

XX (CORR) CORNELL RES FOUND INC.

XX Yoder OC, Turgeon BC, Lu S;

DR WPI: 2001-367672/38.

DR N-PSDB; AAS06685.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

DE Pyrenophora teres partial CPS1 peptide synthetase.
 XX
 KM CPS1; peptide synthetase; peptide toxin; fungal pathogen;
 KM corn crop infection.

OS Pyrenophora teres.

FT Key Location/Qualifiers

FT Region 73..86 /note= "Peptide synthetase motif, core 3"

FT Region 118..125 /note= "Peptide synthetase motif, core 4"

FT Region 135..151 /note= "Peptide synthetase motif, core 5"

FT Region 217..225 /note= "Cyclization domain motif"

FT Region 221..224 /note= "Peptide synthetase motif, core 6"

FT Region 424..435 /note= "Peptide synthetase motif, core 1"

FT Region 524..543 /note= "Peptide synthetase motif, core 2"

FT Region 755..761 /note= "Peptide synthetase motif, core 3"

XX MO200138489-A2.

XX 31-MAY-2001.

XX 22-NOV-2000; 2000MO-US32227.

XX 23-NOV-1999; 99US-0448215.

XX (CORR) CORNELL RES FOUND INC.

XX Yoder OC, Turgeon BC, Lu S;

DR WPI: 2001-367672/38.

DR N-PSDB; AAS06685.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

DE Shrimp white spot Bacilliform virus (WSBV) protein 106.
 XX Shrimp white spot Bacilliform virus; WSBV; diagnosis; viral infection;
 KM antiviral agent; gene expression; antisense construct;
 KM transgenic viral resistant shrimp.
 XX
 OS white spot syndrome virus.
 XX
 PN MO200138351-A2.
 XX
 PD 31-MAY-2001.
 XX
 PF 08-NOV-2000; 2000MO-US28888.
 XX
 PR 24-NOV-1999; 99CN-0124717.
 XX
 PA (PENY-) PE CORP NY.
 PA (THIR-) THIR INST OCEANOGRAPHY STATE OCEANI C A.
 PA (SINO-) SINOENOMAX CO LTD.
 XX
 PI Xu X, Yang F, He J, Pham L, He M, Ye Y, Shen Y, Kodira C;
 XX
 PT WPI: 2001-355877/37.
 DR N-PSDB; AAH62795.
 XX
 XX Primary nucleotide sequence of the shrimp white spot Bacilliform virus
 PT (MSBV), useful for producing viral polypeptides that can be used to
 PT screen for agents that are useful for treating WSBV infection -
 XX
 PS Claim 1: Figure 3: 626pp; English.
 XX
 CC The invention provides the primary nucleotide sequence of the WSBV genome
 CC (AAH62889), predicted transcript sequences (AAH62889-AAH62839) and
 CC encoded proteins (AAH64910-AAH65051) and oligonucleotide sequences
 CC (AAH62840-63160) suitable for use as primers or probes. The nucleic acid
 CC molecules and proteins of the invention are useful for diagnosis and
 CC monitoring viral infection. In screens for antiviral agents and for
 CC monitoring viral gene expression or activity during a treatment regimen.
 CC The nucleic acid molecules are also useful as antisense constructs to
 CC control viral gene expression in infected cells and tissues and to create
 CC transgenic viral resistant shrimp.
 CC
 SO Sequence 1245 AA;
 XX
 Query Match 93.1%; Score 27; DB 22; Length 1245;
 Best Local Similarity 83.3%; Pred. No. 4.1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 REIND 6
 Db 792 relind 797
 XX
 RESULT 7
 ID AA72392 standard; Protein: 259 AA.
 XX
 AC AA72392;
 XX
 DT 24-APR-2001 (first entry)
 XX
 DE Corn LHT1 encoded by cDNA clone ctain.pk0048.h2.
 XX
 XX Corn; lysine- and histidine-specific transporter; LHT1;
 KM proline transporter; transgenic plant; immunological screening.
 XX
 OS Zea mays.
 XX
 PN US6165792-A.
 XX
 PR 16-MAY-1995; 95WO-US06119.
 XX
 PD 26-DEC-2000.
 XX

PF 09-AUG-1999; 99US-0370253.
 XX
 PR 20-AUG-1998; 98US-0097222.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 PI Rhopre CJ, Allen SM;
 XX
 DR WPI: 2001-090485/10.
 DR N-PSDB; AAD02471.
 XX
 XX Nucleic acids encoding amino acid transporters, e.g. lysine- and
 PT histidine-specific transporter and proline transporter, useful for
 PT transgenic plant production with altered expression level of the amino
 PT acid transporters -
 XX
 PS Claim 1: Column 25-26; 24pp; English.
 XX
 CC The present sequence is corn lysine- and histidine-specific transporter
 CC (LHT1) from clone ctain.pk0048.h2, isolated from a corn tissue ctain
 CC cDNA library.
 CC The invention relates to lysine- and histidine-specific transporter
 CC (LHT1), proline transporter and their corresponding cDNA clones. It is
 CC useful for producing transgenic plants with altered expression level of
 CC amino acid transporters, in cell types or developmental stages in which
 CC they are not normally found, such that the obtained transgenic plants
 CC show increased tolerance to salt and water stress. The level of lysine,
 CC histidine and proline in cells of the transgenic plant, is also altered.
 CC The invention is also used for constructing a chimeric gene encoding all
 CC or a portion of the amino acid transporter, where in the expression of
 CC chimeric gene results in the production of altered levels of amino acid
 CC transporter in a transformed host cell. Synthetic peptides obtained from
 CC LHT1 and proline transporter are used for immunising animals to produce
 CC polyclonal or monoclonal antibodies, which facilitates the immunological
 CC screening of cDNA expression libraries.
 CC
 SO Sequence 259 AA;
 XX
 Query Match 89.7%; Score 26; DB 22; Length 259;
 Best Local Similarity 83.3%; Pred. No. 1.4e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 REIND 6
 Db 246 rqlind 251
 XX
 RESULT 8
 ID AAR83032 standard; peptide: 58 AA.
 XX
 AC AAR83032;
 XX
 DT 03-APR-1996 (first entry)
 XX
 DE Polypeptide encoded by 5' flanking region of cps gene.
 XX
 KM Capsular polysaccharide; cps; peptide; flanking region; detection;
 KM serotype; diagnosis; prevention; Streptococcus pneumoniae.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN MO9531548-A1.
 XX
 PD 23-NOV-1995.
 XX
 PF 16-MAY-1995; 95WO-US06119.
 XX
 PR 16-MAY-1994; 94US-0243546.
 XX
 PA (UABR-) UAB RES FOUND.
 XX

PI Dillard J, Yother J;
 XX
 DR WPI: 1996-010934/01.
 DR N-PSDB; AAT05485.
 XX
 PT New Streptococcus pneumoniae capsular polysaccharide genes - used
 for detection, serotyping and for diagnosis and prevention of S.
 pneumoniae infection
 PS
 PS Claim 3; Page 166; 226pp; English.
 CC This polypeptide is encoded by the 5' flanking region of the
 CC capsular polysaccharide gene (cps) of Streptococcus pneumoniae. The
 CC coding sequence is of sufficient length to allow hybridisation under
 CC standard hybridisation conditions to a S.pneumoniae cps gene
 CC flanking region and may be used in methods to detect and serotype S.
 CC pneumoniae. It may also be used for the diagnosis and prevention of
 CC S. pneumoniae infection.
 CC
 SQ Sequence 58 AA;

Query Match 86.2%; Score 25; DB 17; Length 58;
 Best Local Similarity 83.3%; Pred. No. 53;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RETIMD 6
 Db 52 rklind 57

RESULT 9
 AAY33889
 ID AAY33889 standard; Protein: 58 AA.
 XX
 AC AAY33889;
 XX
 DT 04-JUN-2000 (first entry)
 XX
 DE Amino acid sequence of Cps3b.
 XX
 KW Infection; bacterial capsule; serotype specific; virulence;
 KW capsule synthesis; cassette model; binary model;
 KW cassette; in; flanking region; non-serotype specific.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN USS948900-A.
 XX
 PD 07-SEP-1999.
 XX
 PF 02-JUN-1997; 97US-0867030.
 XX
 PR 02-JUN-1997; 97US-0867030.
 PR 16-MAY-1994; 94US-0243546.
 XX
 PA (UABR-) UAB RES FOUND.
 XX
 PI Dillard J, Yother J;
 XX
 DR WPI: 1999-517979/43.
 DR N-PSDB; AAZ11798.
 XX
 PT Non-serotype-specific flanking nucleotide sequences derived from
 PT Streptococcus pneumoniae capsular polysaccharide genes, useful as
 PT hybridization probes for identifying serotype-specific capsular
 PT polysaccharide genes
 PS
 PS Claim 2; Fig 6a; 140pp; English.
 CC This is the amino acid sequence of the capsule synthesis protein
 CC Cps3b.
 CC Serotype-specific cps genes encode the various enzymatic

CC functions of capsule synthesis and hence determine the particular
 CC structure of the capsule polysaccharides produced. Therefore cps genes
 CC determine the serotype of the virus.
 CC The non-serotype specific flanking nucleotide sequences (AAZ11798
 CC AAZ11799, AAZ11800, AAZ11801 and AAZ11802) may be used in hybridization
 CC assays to identify the location of DNA flanking serotype-specific cps
 CC genes in any strain of S. pneumoniae virus.
 CC Additionally, the flanking regions are involved in recombination
 CC and integration of the type specific cps genes during virus
 CC replication. Therefore, when a selected gene (such as the
 CC serotype-specific cps genes) is integrated between the nucleic acids
 CC of the flanking region, the resulting construct can be stably integrated
 CC into an S.pneumoniae host.
 CC This allows the locus of the serotype-specific cps genes (and other
 CC genetic elements), within a virus, to be identified, and therefore,
 CC isolated and characterized. The cps gene sequences may then be used as
 CC hybridization probes to identify and quantify specific serotypes of S.
 CC pneumoniae or to recombinantly produce the gene products for use as
 CC antigens in immunization protocols.
 CC
 SQ Sequence 58 AA;

Query Match 86.2%; Score 25; DB 20; Length 58;
 Best Local Similarity 83.3%; Pred. No. 53;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RETIMD 6
 Db 52 rklind 57

RESULT 10
 AAY75542
 ID AAY75542 standard; Protein: 102 AA.
 XX
 AC AAY75542;
 XX
 DT 21-MAR-2000 (first entry)
 XX
 DE Neisseria meningitidis ORF 747 protein sequence SEQ ID NO:2558.
 XX
 KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
 KW antibacterial; gene therapy.
 XX
 OS Neisseria meningitidis.
 XX
 PN WO9957280-A2.
 XX
 PD 11-NOV-1999.
 XX
 PF 30-APR-1999; 99WO-US09346.
 XX
 PR 01-MAY-1998; 98US-0083758.
 PR 31-JUL-1998; 98US-0094869.
 PR 02-SEP-1998; 98US-0098994.
 PR 02-SEP-1998; 98US-0099062.
 PR 09-OCT-1998; 98US-0103749.
 PR 09-OCT-1998; 98US-0103749.
 PR 09-OCT-1998; 98US-0103796.
 PR 25-FEB-1999; 99US-0121528.
 XX
 PA (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M,
 PI Petersen J, Pizzo M, Rappelli R, Ratti G, Scalato E, Scarselli M,
 PI Tettelin H, Venter JC;
 XX
 DR WPI: 2000-062150/05.
 DR N-PSDB; AAZ54304.
 XX

PT Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics
PS Claim 2; Page 1216; 1453pp; English.
XX
CC AA53015 to AA254536, AA254577 to AA254615, and AA774253 to AA75941
CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
CC and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.
XX
SQ Sequence 102 AA:

Query Match 86.2%; Score 25; DB 21; Length 102;
Best Local Similarity 66.7%; Pred. No. 95;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 REIMD 6
|||:|
Db 68 relvid 73

RESULT 11
AA75543
ID AA75543 standard; Protein; 102 AA.
XX
AC AA75543;
XX
DT 21-MAR-2000 (first entry)
XX
DE Neisseria meningitidis ORF 747 protein sequence SEQ ID NO:2560.
XX
KM Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KM antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
KM antibacterial; gene therapy.
XX
OS Neisseria meningitidis.
XX
PN WO957280-A2.
XX
PD 11-NOV-1999.
XX
PF 30-APR-1999; 99WO-US09346.
XX
PR 01-MAY-1998; 98US-0083758.
PR 31-JUL-1998; 98US-0084869.
PR 02-SEP-1998; 98US-0089894.
PR 02-SEP-1998; 98US-0099062.
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103796.
PR 25-FEB-1999; 99US-0121528.
XX
PA (CHIR) CITRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Fraser C, Galeotti C, Grandi G, Hickey E, Maignani V, Mora M,
PI Petersen J, Pizza M, Rappelli R, Ratti G, Scalato E, Scarselli M,
PI Tettelin H, Venter JC;
XX
XX MPI; 2000-062150/05.
DR N-PSDB; AA254305.
XX
PT Novel Neisserial polypeptides predicted to be useful antigens for

PT vaccines and diagnostics
XX
PS Claim 2; Page 1216; 1453pp; English.
XX
CC AA23015 to AA254536, AA254577 to AA254615, and AA774253 to AA75941
CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
CC and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.
XX
SQ Sequence 102 AA:

Query Match 86.2%; Score 25; DB 21; Length 102;
Best Local Similarity 66.7%; Pred. No. 95;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 REIMD 6
|||:|
Db 68 relvid 73

RESULT 12
AA21412
ID AA21412 standard; Protein; 174 AA.
XX
AC AA21412;
XX
DT 01-APR-1992 (first entry)
XX
DE NADH dehydrogenase 6.
XX
KM Pneumonia; assay; AIDS; immunosuppressed.
XX
OS Pneumocystis carinii.
XX
PN WO9119005-A.
XX
PD 12-DEC-1991.
XX
PF 31-MAY-1991; 91WO-GB00869.
XX
PR 01-JUN-1990; 90GB-0012196.
XX
PA (ISIS-) ISIS INNOVATION LTD.
XX
PA Wakefield AE, Hopkin JM, Moxon ER;
XX
XX MPI; 1992-007487/01.
DR N-PSDB; AAQ20065.
XX
XX New DNA sequences which act as oligo:nucleotide primers - for
PT assaying DNA sample from respiratory secretion of a patient
PT infected with P carinii
XX
PS Claim 3; Fig 3; 42pp; English.
XX
XX The amino acid sequence is that of P. carinii NADH dehydrogenase
CC 6 which was translated from DNA from plasmid pat112. See also
CC AA21409-R21413 and AA20056.
XX
SQ Sequence 174 AA:

Query Match 86.2%; Score 25; DB 13; Length 174;

Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RETIMD 6
Db 111 relime 116

RESULT 13

AA86073
ID AAY86073 standard; Protein; 212 AA.

XX
AC AAY86073;

XX
DT 10-APR-2000 (first entry)

XX
DE S. pneumoniae derived protein #282.

XX
KW Treatment; prevention; disease; diagnosis; gene therapy; screening;
bacterial; antimicrobial; antibiotic; pathogenesis; infection.

XX
OS Streptococcus pneumoniae.

XX
PN MO9806734-A1.

XX
PD 19-FEB-1998.

XX
PF 15-AUG-1997; 97MO-US14436.

XX
PR 16-AUG-1996; 96US-0024022.

XX
PA (SMIK) SMITHKLINE BEECHAM CORP.

XX
PI Black MT, Hodgson JE, Knowles DJC, Lonetto MA, Nicholas RO;

XX
PI Stodola RK;

DR
N-PSDB; AA296394.

XX
PT Streptococcus pneumoniae proteins and related DNA - useful for
screening compounds for antibacterial activity

XX
PS Claim 5; Page 551-552; 640pp; English.

XX
CC This invention describes novel isolated Streptococcus pneumoniae
polynucleotides (see AA296173-296494) and their encoded proteins (see
AA295792-296182). The DNA, vectors and host cells described in the
method of the invention are useful for the recombinant expression of the
polypeptides. The polypeptides are useful for treatment or prevention of
disease, or diagnosis of disease related to expression or activity of
such a polypeptide. They can also be used to screen for compounds which
interact with and inhibit or activate such a polypeptide. The
polypeptides (or DNA encoding them, via gene therapy) are also useful
for inducing an immunological response in a mammal. The antagonists are
useful to inhibit such bacterial polypeptides. The polypeptides are
particularly useful to identify antimicrobial compounds and antibiotics.
CC They are also useful to determine their role in pathogenesis of
infection, dysfunction and disease.

XX
SQ Sequence 212 AA;

Query Match 86.2%; Score 25; DB 19; Length 212;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 RETIMD 6
Db 51 revlid 56

RESULT 14

AA808637

ID AAB08637 standard; Protein; 212 AA.

XX
AC AAB08637;

XX
DT 20-DEC-2000 (first entry)

XX
DE Amino acid sequence of a thymidylate kinase polypeptide.

XX
KW Thymidylate kinase; tdk polypeptide; bacterial infection.

XX
OS Streptococcus pneumoniae.

XX
PN WO200050602-A1.

XX
PD 31-AUG-2000.

XX
PF 18-FEB-2000; 2000MO-US04238.

XX
PR 26-FEB-1999; 99US-0259109.

XX
PA (SMIK) SMITHKLINE BEECHAM CORP.

XX
PI Burnham MKR, Zalacain M, Biswas S, Chalker AF, Ingraham KA;

XX
PI Trelni CM, Warren PV;

DR
N-PSDB; AAA64398.

XX
PF Nucleic acids encoding thymidylate kinase family polypeptides derived
from Streptococcus pneumoniae, useful for screening for antibacterial
agents -

XX
PS Claim 1; Page 3; 41pp; English.

XX
CC The present sequence represents a polypeptide of the thymidylate kinase
family (tdk polypeptides), and is derived from Streptococcus pneumoniae.
CC tdk polynucleotides may be used to recombinantly produce the tdk
polypeptides either in vivo (e.g. as part of a genetic vaccination
procedure) or in vitro (e.g. as part of a fermentation culture).
CC The nucleic acids and proteins may be used to diagnose diseases
in which the tdk polypeptides are expressed, such as infection by
Streptococcus pneumoniae.

XX
SQ Sequence 212 AA;

Query Match 86.2%; Score 25; DB 21; Length 212;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 RETIMD 6
Db 51 revlid 56

RESULT 15

AA81763
ID AAY81763 standard; Protein; 212 AA.

XX
AC AAY81763;

XX
DT 02-JUN-2000 (first entry)

XX
DE Streptococcus pneumoniae protein sequence ID119.

XX
KW Streptococcus pneumoniae infection; immunogen; antigen; diagnosis; AIDS;

XX
KW bacterial pneumonia; asplenia; heart disease; lung disease; alcoholism;

XX
KW kidney disease; diabetes; immunosuppressive disorder; otitis media;

XX
OS Streptococcus pneumoniae.

XX
PN WO200006738-A2.

XX 10-FEB-2000.
 PD
 XX 27-JUL-1999; 99WO-GB02452.
 PF
 XX 27-JUL-1998; 98GB-0016336.
 PR 19-MAR-1999; 99US-0125329.
 XX
 PA (MICR-) MICROBIAL TECHNIQS LTD.
 XX
 PI Le Page RWF, Wells JM, Hanniffy SB, Hansbro PM;
 XX WPI: 2000-195301/17.
 DR N-PSDB: AA291859.
 XX
 PT Streptococcal proteins and polynucleotides useful for diagnosis,
 XX treatment and prophylaxis of bacterial infections -
 XX
 PS Claim 2; Page 61; 76pp; English.
 XX
 CC This sequence represents a Streptococcus pneumoniae protein of the
 CC invention. The proteins (or their homologues, derivatives and/or
 CC fragments) are useful as immunogens or antigens. Immunogenic or antigenic
 CC compositions comprising the proteins are useful as vaccines and also in
 CC diagnostic assays. The sequences are useful for the detection or
 CC diagnosis of S. pneumoniae infection, by contacting a sample to be tested
 CC with them. Agents capable of antagonising, inhibiting or interfering with
 CC the function or expression of the protein or polypeptide are useful in
 CC medical compositions in the treatment or prophylaxis of S. pneumoniae
 CC infection. As the sequences can be used to treat S. pneumoniae infection,
 CC they can be used to treat bacterial pneumonia, which has high rates in
 CC young children, the elderly, and in patients with predisposing conditions
 CC such as asplenia, heart, lung and kidney disease, diabetes, alcoholism,
 CC or with immunosuppressive disorders, especially AIDS. They can also be
 CC used to treat pneumococcal septicaemia, otitis media, sinusitis, and
 CC meningitis.
 CC
 SQ Sequence 212 AA:

Query Match 86.2%; Score 25; DB 21; Length 212;
 Best Local Similarity 66.7%; Pred. No. 2e+02; 0; Indels 0; Gaps 0;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 REITMD 6
 11:1:1
 Db 51 revltd 56

Search completed: March 18, 2002, 09:48:56
 Job time: 365 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:49:41 ; Search time 37.77 Seconds
(without alignments)

3.575 Million cell updates/sec

Title: US-09-544-665-4

Sequence: 1 REIMD 6

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCITUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	89.7	259	4	US-09-370-253-4	Sequence 4, Appl
2	86.2	58	2	US-08-867-030B-7	Sequence 7, Appl
3	86.2	58	5	PCT-US95-06119-7	Sequence 7, Appl
4	86.2	212	4	US-09-259-109-2	Sequence 2, Appl
5	82.8	33	1	US-08-781-020-6	Sequence 6, Appl
6	82.8	33	4	US-09-038-935-6	Sequence 6, Appl
7	82.8	96	1	US-08-565-386-23	Sequence 23, Appl
8	82.8	344	1	US-08-403-866-9	Sequence 9, Appl
9	82.8	379	3	US-08-840-204-3	Sequence 3, Appl
10	82.8	389	3	US-08-650-275-3	Sequence 3, Appl
11	82.8	389	3	US-09-181-318-3	Sequence 3, Appl
12	82.8	390	1	US-08-121-714-6	Sequence 6, Appl
13	82.8	390	1	US-08-477-108A-6	Sequence 6, Appl
14	82.8	390	1	US-08-477-112-6	Sequence 6, Appl
15	82.8	390	5	PCT-US93-08322-6	Sequence 6, Appl
16	82.8	402	1	US-08-315-461-7	Sequence 7, Appl
17	82.8	402	3	US-08-948-997-4	Sequence 4, Appl
18	82.8	402	3	US-08-840-204-2	Sequence 4, Appl
19	82.8	402	4	US-09-348-817A-4	Sequence 4, Appl
20	82.8	402	4	US-09-026-408-3	Sequence 3, Appl
21	82.8	855	2	US-08-482-090-12	Sequence 12, Appl
22	82.8	856	2	US-08-481-700B-8	Sequence 8, Appl
23	82.8	856	2	US-09-007-383-16	Sequence 16, Appl
24	82.8	863	1	US-08-325-547-4	Sequence 4, Appl
25	82.8	1724	1	US-08-325-547-3	Sequence 3, Appl
26	79.3	110	5	PCT-US96-03916-4	Sequence 4, Appl
27	79.3	154	1	US-08-426-627-20	Sequence 20, Appl

28	23	79.3	193	2	US-08-336-031-6	Sequence 6, Appl
29	23	79.3	193	5	PCT-US95-06725-6	Sequence 6, Appl
30	23	79.3	251	2	US-08-766-738-1	Sequence 1, Appl
31	23	79.3	251	2	US-08-766-738-3	Sequence 3, Appl
32	23	79.3	275	2	US-08-578-709-13	Sequence 13, Appl
33	23	79.3	281	2	US-08-360-098A-35	Sequence 55, Appl
34	23	79.3	284	5	PCT-US96-03916-61	Sequence 61, Appl
35	23	79.3	294	5	PCT-US96-03916-71	Sequence 71, Appl
36	23	79.3	411	2	US-08-336-031-2	Sequence 2, Appl
37	23	79.3	411	2	US-08-902-853-7	Sequence 7, Appl
38	23	79.3	411	5	PCT-US95-06725-2	Sequence 2, Appl
39	23	79.3	436	4	US-08-584-760A-67	Sequence 67, Appl
40	23	79.3	440	4	US-08-584-760A-1	Sequence 1, Appl
41	23	79.3	447	4	US-09-370-253-10	Sequence 10, Appl
42	23	79.3	471	4	US-08-961-083-158	Sequence 158, App
43	23	79.3	500	2	US-08-578-709-15	Sequence 15, Appl
44	23	79.3	515	4	US-08-942-012B-32	Sequence 32, Appl
45	23	79.3	606	4	US-09-187-124-2	Sequence 2, Appl

ALIGNMENTS

```

RESULT 1
US-09-370-253-4
; Sequence 4, Application US/09370253
; Patent No. 6165792
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Sakai, Hajime
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: Amino Acid Transporters
; FILE REFERENCE: BB-1200
; CURRENT APPLICATION NUMBER: US/09/370,253
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: 60/097,222
; EARLIER FILING DATE: August 20, 1998
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 4
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Zea mays
; US-09-370-253-4

Query Match      89.7% Score 26; DB 4; Length 259;
Best Local Similarity 83.3%; Pred. No. 47;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB      246 RQIMD 251

RESULT 2
US-08-867-030B-7
; Sequence 7, Application US/08867030B
; Patent No. 5948900
; GENERAL INFORMATION:
; APPLICANT: Kocher et al.
; TITLE OF INVENTION: Streptococcus pneumoniae
; TITLE OF INVENTION: Capsular Polysaccharide Genes and Flanking Regions
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Benjamin Aaron Adler, Ph.D., J.D.
; STREET: 8011 Candle Lane
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch floppy

```

COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft Word for Macintosh
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/867,0308
FILING DATE: June 2, 1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/243,546
FILING DATE: May 16, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Benjamin Aaron Adler, Ph.D., J.D.
REGISTRATION NUMBER: 35,423
REFERENCE/DOCKET NUMBER: D5923
TELECOMMUNICATION INFORMATION:
TELEPHONE: (713) 777-2321
TELEFAX: (713) 777-6908
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 58 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: protein
HYPOTHEICAL: no
ANTI-SENSE: no
FRAGMENT TYPE:
ORIGINAL SOURCE:
IMMEDIATE SOURCE:
POSITION IN GENOME:
US-08-867-0308-7

Query Match 86.2%; Score 25; DB 2; Length 58;
Best Local Similarity 83.3%; Pred. No. 17;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REIMD 6
1:||||
DB 52 KRIIMD 57

RESULT 3
PCT-US95-06119-7
; Sequence 7, Application PC/TUS9506119
; GENERAL INFORMATION:
; APPLICANT: STREPTOCOCCUS PNEUMONIAE CAPSULAR
; TITLE OF INVENTION: POLYSACCHARIDE GENES AND FLANKING REGIONS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P. O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: Patent Release #1.0, Version
; SOFTWARE: #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/06119
; FILING DATE: CONCURRENTLY HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/243,546
; FILING DATE: 16-May-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:

NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: AMCY018P--
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 58 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-06119-7

Query Match 86.2%; Score 25; DB 5; Length 58;
Best Local Similarity 83.3%; Pred. No. 17;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REIMD 6
1:||||
DB 52 KRIIMD 57

RESULT 4
US-09-259-109-2
; Sequence 2, Application US/09259109
; Patent No. 6270762
; GENERAL INFORMATION:
; APPLICANT: Burnham, Martin K. R.
; APPLICANT: Zalacain, Magdalena
; APPLICANT: Biswas, Sanjoy
; APPLICANT: Chalker, Allison F.
; APPLICANT: Ingraham, Karen A.
; APPLICANT: Train, Christopher M.
; APPLICANT: Warren, Patrick V.
; TITLE OF INVENTION: tdk
; FILE REFERENCE: GM10201
; CURRENT APPLICATION NUMBER: US/09/259,109
; CURRENT FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-259-109-2

Query Match 86.2%; Score 25; DB 4; Length 212;
Best Local Similarity 66.7%; Pred. No. 67;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 REIMD 6
1:||||
DB 51 REYIMD 56

RESULT 5
US-08-781-020-6
; Sequence 6, Application US/08781020
; Patent No. 5792749
; GENERAL INFORMATION:
; APPLICANT: Wright, H. Tonle
; TITLE OF INVENTION: METHOD AND COMPOSITION FOR LOWERING LOW
; TITLE OF INVENTION: DENSITY LIPOPROTEIN CHOLESTEROL
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WHITHAM, CURTIS, WHITHAM & MCGINN
; STREET: Reston Int'l. Center, 11800 Sunrise Valley Dr.,
; STREET: Suite 900

CITY: Reston
STATE: VA
COUNTRY: USA
ZIP: 20191
COMPUTER READABLE FORM:
MEDIUM TYPE: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA: #1.0, Version #1.30
APPLICATION NUMBER: US/08/781,020
FILING DATE: 09-JAN-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Whitlam, Michael E.
REGISTRATION NUMBER: 32,635
REFERENCE/DOCKET NUMBER: VCU-16
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-391-2510
TELEFAX: 703-391-9035
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-781-020-6

Query Match 82.8%; Score 24; DB 1; Length 33;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 EIIIMD 6
|||||
Db 5 EIIIMD 9

RESULT 6
US-09-038-935-6
Sequence 6, Application US/09038935
Patent No. 6130332
GENERAL INFORMATION:
APPLICANT: Wright, H. Tonie
APPLICANT: Janciuskiene, Sabina
TITLE OF INVENTION: LIPOPROTEIN CHOLESTEROL
FILE REFERENCE: 294022Da
CURRENT APPLICATION NUMBER: US/09/038,935
CURRENT FILING DATE: 1998-03-12
EARLIER APPLICATION NUMBER: 08/781,020
EARLIER FILING DATE: 1997-01-19
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 6
LENGTH: 33
TYPE: PRT
ORGANISM: Homo sapiens
US-09-038-935-6

Query Match 82.8%; Score 24; DB 4; Length 33;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 EIIIMD 6
|||||
Db 5 EIIIMD 9

RESULT 7
US-08-565-386-23

Sequence 23, Application US/08565386
Patent No. 5741697
GENERAL INFORMATION:
APPLICANT: Bayoll, Patrick M.
APPLICANT: Hala, Ru-Ching
TITLE OF INVENTION: BACTERIOPHAGE OF CHLAMYDIA PSITTACI
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
STREET: Clinton Square, P.O. Box 1051
CITY: Roches
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/565,386
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Timian, Susan J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 176/60040
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1600
TELEFAX: 716-263-1636
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 96 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-565-386-23

Query Match 82.8%; Score 24; DB 1; Length 96;
Best Local Similarity 50.0%; Pred. No. 50;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIIIMD 6
|||||
Db 35 RDVYMD 40

RESULT 8
US-08-403-866-9
Sequence 9, Application US/08403866
Patent No. 5643779
GENERAL INFORMATION:
APPLICANT: Ehrlich, Stanislaw
APPLICANT: Godoy, Jean-Jacques
TITLE OF INVENTION: Nucleic acid coding for an alpha-acetolactate
synthase from Lactococcus and its applications
NUMBER OF SEQUENCES: 16
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/403,866
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 20747/30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1600
TELEFAX: (716) 263-1487

TELEX: 978450 (MUT)
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 344 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ORGANISM: Lactococcus lactis subsp. lactis
; INDIVIDUAL ISOLATE: ILVC
US-08-403-866-9

Query Match 82.8%; Score 24; DB 1; Length 344;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 REIMD 6
Db 162 REIMD 167

RESULT 9
US-08-840-204-3
; Sequence 3, Application US/08840204
; Patent No. 6103498
; GENERAL INFORMATION:
; APPLICANT: LAWRENCE, DANIEL A.
; APPLICANT: STEFANSSON, STEINGRIMUR P.
; TITLE OF INVENTION: MUTANT PLASMINOGEN ACTIVATOR-INHIBITOR
; TITLE OF INVENTION: TYPE 1 (P1-1) AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 PENNSYLVANIA AVENUE, NW
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/840,204
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: LYNNAT, SHMUEL
; REGISTRATION NUMBER: 33,949
; REFERENCE/DOCKET NUMBER: 30807-20004.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 822-0168
; TELEX: 90-4030 MRSNFOERSMWH
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-840-204-3

Query Match 82.8%; Score 24; DB 3; Length 379;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 REIMD 6
Db 351 REIMD 355

RESULT 10
US-08-650-275-3
; Sequence 3, Application US/08650275
; Patent No. 5798249
; GENERAL INFORMATION:
; APPLICANT: Braxton, Scott Michael
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: HUMAN PROTEIN DISULFIDE ISOMERASE
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/650,275
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Luther, Barbara J.
; REGISTRATION NUMBER: 33,954
; REFERENCE/DOCKET NUMBER: PF-0067 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-852-0195
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 389 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: GI 1086627
US-08-650-275-3

Query Match 82.8%; Score 24; DB 1; Length 389;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 REIMD 6
Db 313 REIMD 318

RESULT 11
US-09-181-318-3
; Sequence 3, Application US/09181318
; Patent No. 6001632
; GENERAL INFORMATION:
; APPLICANT: Braxton, Scott Michael
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: HUMAN PROTEIN DISULFIDE ISOMERASE
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette

```

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/181,318
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/650,275
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Luther, Barbara J.
REGISTRATION NUMBER: 33,954
REFERENCE/DOCKET NUMBER: PF-0067 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-852-0195
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 389 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: GI 1086627
US-09-181-318-3

```

```

Query Match      82.8%; Score 24; DB 3; Length 389;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 REIMD 6
DB 313 REFWD 318

```

```

RESULT 12
US-08-121-714-6
Sequence 6, Application US/08121714
Patent No. 5470970
GENERAL INFORMATION:
APPLICANT: Sager, Ruth
TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH
NUMBER OF SEQUENCES: 8
TUMOR SUPPRESSING ACTIVITY
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
OPERATING SYSTEM: IBM PS/2 Model 502 or 55SX
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/121,714
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/938,823
FILING DATE: 09/01/92
APPLICATION NUMBER: 07/844,296
FILING DATE: 02/28/92
APPLICATION NUMBER: 07/662,216
FILING DATE: 02/28/91
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819

```

```

REFERENCE/DOCKET NUMBER: 00530/072001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 390
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-121-714-6

```

```

Query Match      82.8%; Score 24; DB 1; Length 390;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 EIMD 6
DB 374 EIMD 378

```

```

RESULT 13
US-08-477-108A-6
Sequence 6, Application US/08477108A
Patent No. 5801001
GENERAL INFORMATION:
APPLICANT: Sager, Ruth
APPLICANT: Zou, Zhigang
APPLICANT: Anisowicz, Anthony
TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH
NUMBER OF SEQUENCES: 11
TUMOR SUPPRESSING ACTIVITY
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
OPERATING SYSTEM: IBM PS/2 Model 502 or 55SX
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,108A
FILING DATE: June 7, 1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/121,714
FILING DATE: 09/01/93
APPLICATION NUMBER: 07/938,823
FILING DATE: 09/01/92
APPLICATION NUMBER: 07/844,296
FILING DATE: 02/28/92
APPLICATION NUMBER: 07/662,216
FILING DATE: 02/28/91
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 06570/002002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 390
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear

```

US-08-477-108A-6

Query Match 82.8%; Score 24; DB 1; Length 390;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EIMD 6
|||||
DB 374 EIMD 378

RESULT 14
US-08-477-112-6
; Sequence 6, Application US/08477112
; Patent No. 5905023
; GENERAL INFORMATION:
; APPLICANT: Sager, Ruth
; TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH
; TUMOR SUPPRESSING ACTIVITY
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: Wordperfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,112
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/121,714
; FILING DATE: 09/01/93
; APPLICATION NUMBER: 07/938,823
; FILING DATE: 09/01/92
; APPLICATION NUMBER: 07/844,296
; FILING DATE: 02/28/92
; APPLICATION NUMBER: 07/662,216
; FILING DATE: 02/28/91
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 06570/002003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 390
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-477-112-6

Query Match 82.8%; Score 24; DB 2; Length 390;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EIMD 6
|||||
DB 374 EIMD 378

RESULT 15

PCT-US93-08322-6

; Sequence 6, Application PC/TUS9308322
; GENERAL INFORMATION:
; APPLICANT: Sager, Ruth
; TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH TUMOR SUPPRESSING ACTIVITY
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: Wordperfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08322
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/938,823
; FILING DATE: 09/01/92
; APPLICATION NUMBER: 07/844,296
; FILING DATE: 02/28/92
; APPLICATION NUMBER: 07/662,216
; FILING DATE: 02/28/91
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/072001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 390
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
PCT-US93-08322-6

Query Match 82.8%; Score 24; DB 5; Length 390;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EIMD 6
|||||
DB 374 EIMD 378

Search completed: March 18, 2002, 09:49:41
Job time: 225 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Comphen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:50:31 ; Search time 41.98 Seconds

(without alignments)
10.887 Million cell updates/sec

Title: US-09-544-665-4

Sequence: 1 REIMD 6

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	DB ID	Description
1	29	100.0	380 2 C64364	formate hydrogently
2	29	100.0	413 2 T31823	hypothetical prote
3	28	96.6	456 2 B70482	oxygen-independent
4	28	96.6	817 2 F86742	ribonuclease (lipo
5	27	93.1	243 2 G71209	hypothetical prote
6	27	93.1	340 2 D64703	branched-chain-am
7	27	93.1	437 2 A69172	conserved hypothet
8	27	93.1	590 2 B75377	glucose-6-phosphat
9	26	89.7	229 2 A71633	DNA polymerase III
10	26	89.7	306 2 S05658	band 3 anion trans
11	26	89.7	326 2 A71681	pyruvate dehydroge
12	26	89.7	327 2 B71661	hypothetical prote
13	26	89.7	341 2 S52527	hypothetical prote
14	26	89.7	384 2 A84955	cell division prot
15	26	89.7	476 2 T41879	ACNMPV orf142 - Bo
16	26	89.7	477 2 H72867	early 49 kDa prote
17	26	89.7	580 2 T45064	arginine deiminase
18	26	89.7	651 2 E82299	cell division prot
19	26	89.7	747 2 T40728	hypothetical prote
20	26	89.7	973 2 U00971	transposase tmpr -
21	26	89.7	1390 2 T14004	trfA protein - sll
22	25	86.2	110 2 T08921	hypothetical prote
23	25	86.2	146 2 D84083	nitrogen fixation
24	25	86.2	147 1 E70019	iron-sulfur cofact
25	25	86.2	175 2 T31257	terminal oxygenase
26	25	86.2	201 2 T50933	dioxysenem-1 PSI-
27	25	86.2	235 2 S51813	hypothetical prote
28	25	86.2	242 2 S19700	hypothetical prote
29	25	86.2	243 2 T50031	capsular polysacch

ALIGNMENTS

30	25	86.2	278 2 S51774	hypothetical prote
31	25	86.2	292 2 C84865	hypothetical prote
32	25	86.2	309 2 B71195	hypothetical prote
33	25	86.2	313 2 F70333	hypothetical prote
34	25	86.2	346 2 D84833	hypothetical prote
35	25	86.2	351 2 T41572	hypothetical prote
36	25	86.2	366 2 T46249	formate hydrogently
37	25	86.2	370 2 F69151	protein T23G18.18
38	25	86.2	382 2 A86217	hypothetical prote
39	25	86.2	386 2 C83065	hypothetical prote
40	25	86.2	391 1 F70477	conserved hypothet
41	25	86.2	480 2 C86174	hypothetical prote
42	25	86.2	495 2 T31203	hypothetical prote
43	25	86.2	525 2 S19701	hypothetical prote
44	25	86.2	531 2 G84710	hypothetical prote
45	25	86.2	682 2 B84415	translation elonga

RESULT 1
C64364
formate hydrogentlyase, subunit 5 - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C:Accession: C64364
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blum, R.; Bult, C.J.; Overbeek, R.; Kirkness, E.F.; Weissbach, K.G.; Merrick, J.M.; Glodek, I.; Ren, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurest, M.A.; Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woel, A.; Riddle: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A:Reference number: A84300; MUID:96357999
A:Accession: C64364
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-380 <BUL>
A:Cross-references: GB:U67501; GB:L77117; NID:g2826289; PIDN:AAB98504.1; PID:g15912;
C:Genetics:
A:Map position: REV458767-457625
C:Superfamily: NMDH dehydrogenase (ubiquinone) 49k protein

Query Match 100.0%; Score 29; DB 2; Length 380;
Best local similarity 100.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 REIMD 6
DB 133 REIMD 138

RESULT 2
T31823
hypothetical protein C17E7.8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T31823
R:Bradshaw, H.
submitted to the EMBL Data Library, July 1997
A:Description: The sequence of C. elegans consd C17E7.
A:Reference number: Z21091
A:Accession: T31823
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-413 <BRA>
A:Cross-references: EMBL:AF016443; PIDN:AC24282.1; GSPDB:GNO0023; CESP:C17E7.8
A:Experimental source: strain Bristol N2; clone C17E7
C:Genetics:
A:Gene: CESP:C17E7.8
A:Map position: 5
A:Introns: 27/3; 64/1; 164/3; 242/3; 326/1

Query Match 100.0%; Score 29; DB 2; Length 413;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RETIMD 6
 |||||
 DB 290 RETIMD 295

RESULT 3

oxygen-independent coproporphyrinogen II - Aquifex aeolicus
 B70482
 C:Species: Aquifex aeolicus
 C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 11-Jun-1999
 C:Accession: B70482
 R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V.
 Nature 392, 353-358, 1998
 A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
 A:Reference number: A70300; MUID:98196566
 A:Accession: B70482
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-456 <AOF>
 A:Cross-references: GB:AE000774; NID:92984324; PIDN:AA07846.1; PID:92984325; GB:AE00065
 A:Experimental source: strain VP5
 C:Genetics:
 A:Gene: hemA
 C:Superfamily: oxygen-independent coproporphyrinogen oxidase

Query Match 96.6%; Score 28; DB 2; Length 456;
 Best Local Similarity 83.3%; Pred. No. 36;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RETIMD 6
 |||||
 DB 371 RETIMD 376

RESULT 4

ribonuclease [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
 F86742
 C:Species: Lactococcus lactis subsp. lactis
 C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 31-Mar-2001
 C:Accession: F86742
 R:Bohlooli, A.; Winkler, P.; Manger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Ehrlich
 Genome Res. in press, 2001
 A:Title: The complete genome sequence of the lactic acid bacterium.
 A:Reference number: A86625
 A:Accession: F86742
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-817 <STO>
 A:Cross-references: GB:AE005176; NID:912723878; PIDN:AAK05040.1; GSPDB:GN00146
 A:Experimental source: strain IL1403
 C:Genetics:
 A:Gene: vacB1
 C:Superfamily: virulence-associated protein vacB homolog

Query Match 96.6%; Score 28; DB 2; Length 817;
 Best Local Similarity 83.3%; Pred. No. 69;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RETIMD 6
 |||||
 DB 4 RETIMD 9

RESULT 5

G71209
 hypothetical protein PH1943 - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii
 C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000
 C:Accession: G71209
 R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; S.
 M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Ogi
 DNA Res. 5, 55-76, 1998
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
 A:Reference number: A71000; MUID:98344137
 A:Accession: G71209
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-243 <KAW>
 A:Cross-references: GB:AP000007; NID:9326134; PIDN:BA31070.1; PID:93258387
 A:Experimental source: strain OT3
 A:Note: this accession replaces an interim accession for a sequence replaced by GenB.
 C:Genetics:
 A:Gene: PH1943
 C:Superfamily: Methanococcus conserved hypothetical protein MJ1638

Query Match 93.1%; Score 27; DB 2; Length 243;
 Best Local Similarity 83.3%; Pred. No. 32;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RETIMD 6
 |||||
 DB 234 RETIMD 239

RESULT 6

branched-chain-amino-acid aminotransferase - Helicobacter pylori (strain 26695)
 D64703
 C:Species: Helicobacter pylori
 C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 29-Sep-1999
 C:Accession: D64703
 R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann,
 Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khair, H.G.; Glodek, A.; McK
 son, J.D.; Kelley, J.M.; Cotton, M.D.; Meldrum, J.M.; Fujii, C.; Bowman, C.; Watney
 Nature 388, 539-547, 1997
 A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpe, P.D.; Smith, H.O.; Fraser
 A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
 A:Reference number: A64520; MUID:97394467
 A:Accession: D64703
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-340 <TOM>
 A:Cross-references: GB:AE000647; GB:AE000511; NID:92314645; PIDN:AMD08509.1; PID:923
 C:Superfamily: branched-chain-amino-acid transaminase BMT1

Query Match 93.1%; Score 27; DB 2; Length 340;
 Best Local Similarity 83.3%; Pred. No. 46;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RETIMD 6
 |||||
 DB 270 RETIMD 275

RESULT 7

conserved hypothetical protein MTH545 - Methanobacterium thermoautotrophicum (strain
 A69172
 C:Species: Methanobacterium thermoautotrophicum
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
 C:Accession: A69172
 R:Smith, D.R.; Doucette-Stamm, L.A.; Delonghery, C.; Lee, H.; Dubois, J.; Aldredge,
 ; Qiu, D.; Spedafora, R.; Vicalre, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jivani
 K.L.; Church, G.M.; Daniels, C.J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: f
 A:Reference number: A69000; MUID:98037514

RESULT 12

hypothetical protein RP564 - Rickettsia prowazekii
 B71661
 C:Species: Rickettsia prowazekii
 C:Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
 C:Accession: B71661
 R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichteritz-Ponten, T.; Alismark, U.
 Nature 396, 133-140, 1998
 A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
 A:Reference number: A71630; MUID:99039499
 A:Accession: B71661
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-327 <AND>
 A:Cross-references: GB:AJ235272; GB:AJ235269; NID:g3861033; PIDN:CAAI5012.1; PID:e134285
 A:Experimental source: strain Madrid E
 C:Genetics:
 A:Gene: RP564

Query Match

Best Local Similarity 89.7%; Score 26; DB 2; Length 327;
 Best Local Similarity 83.3%; Pred. No. 80;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RETIMD 6
 1-11111
 DB 153 RETIMD 158

RESULT 13

S52527
 hypothetical protein YPL004c - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein LPA13c; hypothetical protein YPL132.09c
 C:Species: Saccharomyces cerevisiae
 C:Date: 08-May-1995 #sequence_revision 21-Jul-1995 #text_change 12-Dec-1997
 C:Accession: S52527; S59689
 R:Badcock, K.; Churcher, C.
 submitted to the EMBL Data Library, February 1995
 A:Reference number: S52519
 A:Accession: S52527
 A:Molecule type: DNA
 A:Residues: 1-341 <BAD>
 A:Cross-references: EMBL:Z48483; NID:g683777; PID:g683786; MIPS:YPL004c
 A:Experimental source: strain AB972
 R:Hall, J.; Ahmed, A.; Bussey, H.; Fortin, N.; Friesen, J.D.; Storms, R.K.; Vo, D.H.; We
 submitted to the EMBL Data Library, August 1995
 A:Description: The sequence of Saccharomyces cerevisiae chromosome XVI left arm.
 A:Reference number: S59677
 A:Accession: S59689
 A:Molecule type: DNA
 A:Residues: 1-341 <HAL>
 A:Cross-references: EMBL:U33335; NID:g965076; PID:g965089; MIPS:YPL004c
 C:Genetics:
 A:Map position: 16L

Query Match

Best Local Similarity 89.7%; Score 26; DB 2; Length 341;
 Best Local Similarity 83.3%; Pred. No. 83;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RETIMD 6
 1-11111
 DB 248 RETIMD 253

RESULT 14

AB4955
 cell division protein fts2 (imported) - Buchnera sp. (strain APS)
 C:Species: Buchnera sp.
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
 C:Accession: AB4955
 #strains: H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.

Nature 407, 81-86, 2000
 A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp.
 A:Reference number: AB4930; MUID:20445173
 A:Accession: AB4955
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-384 <STO>
 A:Cross-references: GB:AF000398; GSPDB:GN00144
 A:Experimental source: strain APS
 C:Genetics:
 A:Gene: fts2; BU212
 C:Superfamily: cell division protein fts2

Query Match

Best Local Similarity 89.7%; Score 26; DB 2; Length 384;
 Best Local Similarity 66.7%; Pred. No. 93;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 RETIMD 6
 1-11111
 DB 332 RETIMD 337

RESULT 15

T41879
 ACMPV orf142 - Bombyx mori nuclear polyhedrosis virus (isolate T3)
 C:Species: Bombyx mori nuclear polyhedrosis virus, BMSNPV
 A:Variety: isolate T3
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
 C:Accession: T41879
 R:Gomi, S.; Majima, K.; Maeda, S.
 J. Gen. Virol. 80, 1323-1337, 1999
 A:Title: Sequence analysis of the genome of Bombyx mori nucleopolyhedrovirus.
 A:Reference number: 222020; MUID:99281911
 A:Accession: T41879
 A:Status: preliminary; translated from GB/EMBL/DDBB
 A:Molecule type: DNA
 A:Residues: 1-476 <KAN>
 A:Cross-references: EMBL:L33180; NID:g3745835; PIDN:AAC63808.1; PID:g3745961
 A:Experimental source: isolate T3
 C:Genetics:
 A:Note: Orf.118
 C:Superfamily: Orgyia pseudotsugata nuclear polyhedrosis virus hypothetical protein 1

Query Match 89.7%; Score 26; DB 2; Length 476;
 Best Local Similarity 83.3%; Pred. No. 1,2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RETIMD 6
 1-11111
 DB 320 RETIMD 325

Search completed: March 18, 2002, 09:50:32
 Job time: 241 sec

A:Accession: A69172
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-437 <MTH>
A:Cross-references: GB:AE000837; GB:AE000666; NID:g2621613; PIDN:AAB85051.1; PID:g262162
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH545

Query Match 93.1%; Score 27; DB 2; Length 437;
Best Local Similarity 83.3%; Pred. No. 62;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REIMD 6
||||:|
Db 63 REIMD 68

RESULT 8
B75377

glucose-6-phosphate 1-dehydrogenase - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000

C:Accession: B75377

R:White, O.; Eisen, J.A.; Heldelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

S.; Shen, M.; Vamathevan, J.; Lam, P.; McDonald, L.; Uitterback, T.; Zalewski, C.; Ma

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896

A:Accession: B75377

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-590 <MTH>

A:Cross-references: GB:AE002003; GB:AE000513; NID:g6459358; PIDN:AAF1158.1; PID:g645936

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR1596

A:Map position: 1

C:Superfamily: glucose-6-phosphate dehydrogenase

Query Match 93.1%; Score 27; DB 2; Length 590;
Best Local Similarity 66.7%; Pred. No. 86;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 REIMD 6
||||:|
Db 500 REVMD 505

RESULT 9
A71633

DNA polymerase III, epsilon chain (dna) RP732 - Rickettsia prowazekii

C:Species: Rickettsia prowazekii

C>Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000

C:Accession: A71633

R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichteritz-Ponten, T.; Alsmark, U

Nature 396, 133-140, 1998

A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.

A:Reference number: A71630; MUID:99039499

A:Accession: A71633

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-229 <AND>

A:Cross-references: GB:AJ235273; GB:AJ235269; NID:g3861237; PIDN:CAA15161.1; PID:g386126

A:Experimental source: strain Madrid E

C:Genetics:

A:Gene: pphA; RP261

C:Superfamily: pyruvate dehydrogenase (lipoamide) alpha chain; thiamin pyrophosphate

Query Match 89.7%; Score 26; DB 2; Length 229;
Best Local Similarity 83.3%; Pred. No. 53;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REIMD 6
||||:|
Db 5 REIMD 10

RESULT 10
S05658

band 3 anion transport protein - bovine (fragments)

C:Species: Bos primigenius taurus (cattle)

C>Date: 21-Nov-1993 #sequence_revision 28-Apr-1995 #text_change 28-Apr-1995

C:Accession: S05658; A61591

R:Moriyama, R.; Kawamatsu, S.; Kondo, Y.; Tomida, M.; Makino, S.

Arch. Biochem. Biophys. 274, 130-137, 1999

A:Title: Antigenic determinants of the cytoplasmic domain of band 3 from bovine eryt.

A:Reference number: S05658; MUID:89372925

A:Accession: S05658

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-306 <MOR>

R:Moriyama, R.; Makino, S.

Arch. Biochem. Biophys. 256, 606-617, 1987

A:Title: Interaction of glyceraldehyde-3-phosphate dehydrogenase with the cytoplasmic

I band 3 polypeptide.

A:Reference number: A61591; MUID:87297550

A:Accession: A61591

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-64, 'X', 66, 'X', 68-69 <MO2>

C:Superfamily: band 3 anion transport protein

Query Match 89.7%; Score 26; DB 2; Length 306;
Best Local Similarity 66.7%; Pred. No. 74;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 REIMD 6
||||:|
Db 49 REIMD 54

RESULT 11
A71681

pyruvate dehydrogenase E1 component, alpha chain precursor (pdh) RP261 - Rickettsia

C:Species: Rickettsia prowazekii

C>Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000

C:Accession: A71681

R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichteritz-Ponten, T.; Alsmar

Nature 396, 133-140, 1998

A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria

A:Reference number: A71630; MUID:99039499

A:Accession: A71681

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-326 <AND>

A:Cross-references: GB:AJ235271; GB:AJ235269; NID:g3868717; PIDN:CAA14723.1; PID:g386

A:Experimental source: strain Madrid E

C:Genetics:

A:Gene: pdhA; RP261

C:Superfamily: pyruvate dehydrogenase (lipoamide) alpha chain; thiamin pyrophosphate

Query Match 89.7%; Score 26; DB 2; Length 326;
Best Local Similarity 83.3%; Pred. No. 79;
Matches 5; Conservative 1; Mismatches 0; Indels

QY 1 REIMD 6
||||:|
Db 277 REIMD 282

The Mar 19 11:00:06

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:52:26 ; Search time 24.53 Seconds
(without alignments)
8,968 Million cell updates/sec

Title: US-09-544-665-4
Perfect score: 29
Sequence: 1 REIIMD 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	96.6	456	1	HEMN_AQUAE
2	27	93.1	340	1	HEMN_HELPY
3	27	93.1	461	1	TBC_NENCY
4	26	89.7	229	1	DP3E_RICPR
5	26	89.7	326	1	DDPA_RICPR
6	26	89.7	384	1	FTS2_BUCAL
7	26	89.7	404	1	NDD2_RHIME
8	26	89.7	477	1	Y142_NPVAC
9	25	86.2	147	1	NIFU_BACSU
10	25	86.2	235	1	PSAF_HORVU
11	25	86.2	309	1	THIL_PYRHO
12	25	86.2	313	1	X376_AQUAE
13	25	86.2	578	1	FYVJ_HUMAN
14	25	86.2	728	1	EP2_HALHA
15	25	86.2	941	1	MSH1_SCHPO
16	25	86.2	1522	1	BA13_HUMAN
17	25	86.2	1986	1	WALENT
18	25	86.2	96	1	Y11K_BPCHP
19	24	82.8	120	1	CYB5_YEAST
20	24	82.8	123	1	C59A_MOUSE
21	24	82.8	155	1	RL13_AERPE
22	24	82.8	178	1	RL5_ARCFU
23	24	82.8	197	1	GRPE_ECOLI
24	24	82.8	206	1	KTHY_DEIRA
25	24	82.8	270	1	PEGS_ZYMOO
26	24	82.8	292	1	YE95_METJA
27	24	82.8	299	1	BAH_STRHY
28	24	82.8	300	1	NARE_CHICK
29	24	82.8	309	1	P2A_BRANA
30	24	82.8	313	1	NOD3_RHIME
31	24	82.8	313	1	P2A3_ARATH
32	24	82.8	313	1	P2A4_ARATH
33	24	82.8	313	1	P2A_MEDSA

34	24	82.8	337	1	2265_HUMAN	005218	homo sapien
35	24	82.8	340	1	ILVC_LACIA	002138	lactococcus
36	24	82.8	372	1	BHBS_BOVIN	P14893	b 3 beta-hy
37	24	82.8	384	1	FTS2_BUCAP	051929	buchnera ap
38	24	82.8	400	1	PAT1_MUSVI	P50449	musstela vis
39	24	82.8	402	1	PAT1_BOVIN	P13909	bos taurus
40	24	82.8	402	1	PAT1_HUMAN	P05121	homo sapien
41	24	82.8	402	1	PAT1_PIG	P79335	sus scrofa
42	24	82.8	444	1	P44_PANTR	P27473	pan troglod
43	24	82.8	493	1	YDAK_CAEEL	P60771	caenorhabd
44	24	82.8	499	1	C771_SOJME	P71231	solanum mel
45	24	82.8	531	1	TCPY_MOUSE	061390	mus musculus

ALIGNMENTS

RESULT 1
HEMN_AQUAE STANDARD: PRT: 456 AA.
ID HEMN_AQUAE
AC 067886;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE OXYGEN-INDEPENDENT COPROPHORHYRINOGEN II.
GN HEMN OR AQ.2124.
GS Aquifex aeolicus.
OS Bacteria; Aquificales; Aquificaceae; Aquifex.
OC NCBI_TaxID=63363;
RX STRAIN=VF5;
RC MEDLINE=96196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Shead W.A., Keller M., Anjaly M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus.";
RL Nature 392:353-358(1998).
CC -1- FUNCTION: ANAEROBIC TRANSFORMATION OF COPROPHORHYRINOGEN-III INTO
COPROPHORHYRINOGEN-IX (BY SIMILARITY).
CC -1- COPROPHOR: REQUIRES MAGNESIUM, ATP AND NAD (OR NADP) FOR ACTIVITY.
CC -1- PATHWAY: PORPHYRIN BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE ANAEROBIC COPROPHORHYRINOGEN III
OXIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb.ch/announce/>
or send an email to license@isb.slb.ch).
CC -----
DR EMBL: AE000774; AAC07846.1; -
DR InterPro: IPR003401; Coprogen_an_ox.
DR Pfam: PF02473; Coprogen_an_ox; 1.
DR ProPrint: biosynthes; Oxidoreductase; Magnesium; NAD;
KW Complete proteome.
SQ SEQUENCE 456 AA; 53785 MW; 4E494103B8EF6342 CRC64;

Query Match 96.6%; Score 28; DB 1; Length 456;
Best local similarity 83.3%; Pred. No. 19;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REIIMD 6
DB 371 REIIMD 376
RESULT 2

```

LIVE_HELPY
ID LIVE_HELPY STANDARD: PRT; 340 AA.
AC 026004:
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE BRANCHED-CHAIN AMINO ACID AMINOTRANSFERASE (EC 2.6.1.42) (BCAT).
GN LIVE OR HP1468.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-P., White O., Kervatage A., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klein H.-P., Gill S., Dougherty B.A.,
RA Lettun K., Richardson D., Zhou L., Kirness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khair H.G., Glodek A.,
RA McInerney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Meldrum J.M., Fujii C., Bowman C., Matheny L., Mallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RL Nature 388:539-547(1997).
CC -1- CATALYTIC ACTIVITY: L-LEUCINE + 2-OXOGLOUTARATE -> 4-METHYL-2-
CC OXOPENTANOATE + L-GLUTAMATE (ALSO ACTS ON L-ISOLEUCINE AND
CC L-VALINE).
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- PATHWAY: VALINE AND ISOLEUCINE BIOSYNTHESIS.
CC -1- SIMILARITY: BELONGS TO CLASS-IV OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000647; AAD08509.1;
DR HSSP: P00510; 1A3G.
DR TIGR: HP1468;
DR InterPro: IPR001544; AminoTran_4.
DR Pfam: PF01063; aminoTran_4; 1.
DR ProDom: PD001961; AminoTran_4; 1.
DR PROSITE: PS00770; AA_TRANSFR_CLASS_4; 1.
KM Transferrase; Amino transferase; Branched-chain amino acid biosynthesis;
FM Pyridoxal phosphate; Complete proteome.
FT BINDING 187 187 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 340 AA; 37641 MW; 84F1685F61DC52C CRC64;

Query Match 93.1%; Score 27; DB 1; Length 340;
Best Local Similarity 83.3%; Pred. No. 24;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 REIMD 6
DB 270 REIMD 275

RESULT 3
ID TBG_NEUCR STANDARD: PRT; 461 AA.
AC P53377;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)

```

```

DE TUBULIN GAMMA CHAIN (GAMMA TUBULIN).
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98019102; PubMed=9358070;
RA Heckmann S., Schliwa M., Kube-Grandenath E.;
RA Primary structure of Neurospora crassa gamma-tubulin."
RL Gene 199:303-309(1997).
CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES.
CC GAMMA TUBULIN IS FOUND AT MICROTUBULE ORGANIZING CENTERS (MTOC)
CC SUCH AS THE SPINDLE POLES OR THE CENTROSOME, SUGGESTING THAT IT
CC IS INVOLVED IN THE MINUS-END NUCLEATION OF MICROTUBULE ASSEMBLY.
CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X97753; CA66348.1;
DR InterPro: IPR000217; Tubulin.
DR InterPro: IPR003008; Tubulin_FtsZ.
DR Pfam: PF00091; tubulin; 1.
DR PRINTS: PRO1161; TUBULIN.
DR PROSITE: PRO1164; GAMMAYTUBULIN.
DR PROSITE: PS00227; TUBULIN; 1.
KM Microtubules; GTP-binding.
FT BIND 142 148 GTP (POTENTIAL).
SQ SEQUENCE 461 AA; 51598 MW; 70DCBEA224B6343 CRC64;

Query Match 93.1%; Score 27; DB 1; Length 461;
Best Local Similarity 66.7%; Pred. No. 34;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 REIMD 6
DB 425 REIMD 430

RESULT 4
ID DP3E_RICPR STANDARD: PRT; 229 AA.
AC 092CJ9;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE DNA POLYMERASE III, EPSILON CHAIN (EC 2.7.7.7).
GN DNAO OR RP732.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiinae; Rickettsia.
OX NCBI_TaxID=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MADRID E.
RX MEDLINE=99039499; PubMed=9823893;
RA Anderson S.G.E., Zomorodipour A., Anderson J.O.,
RA Scharitz-Ponten T., Almark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA The genome sequence of Rickettsia prowazekii and the origin of
RA mitochondria."
RL Nature 396:133-140(1998).
CC -1- FUNCTION: DNA POLYMERASE III IS A COMPLEX, MULTISUBUNIT ENZYME
CC RESPONSIBLE FOR MOST OF THE REPLICATIVE SYNTHESIS IN BACTERIA.
CC THE EPSILON SUBUNIT CONTAINS THE EDITING FUNCTION AND IS A
CC PROOFREADING 3'-5' EXONUCLEASE (BY SIMILARITY).

```

```

CC      -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE -
CC      N PYROPHOSPHATE + DNA(N).
CC      -1- SUBUNIT: CONTAINS A CORE (COMPOSED OF ALPHA, EPSILON, AND THETA
CC      CHAINS) THAT ASSOCIATES WITH A TAU SUBUNIT WHICH ALLOWS THE CORE
CC      DIMERIZATION TO FORM THE POLIIT' COMPLEX. POLIIT' ASSOCIATES WITH
CC      THE GAMMA COMPLEX (COMPOSED OF CHAINS GAMMA, DELTA, DELTA', PSI,
CC      AND CHI) AND WITH THE BETA CHAIN (BY SIMILARITY).
CC      -----
CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: AJ235273; CAA15161.1; -.
CC      InterPro: IPR000520; Exonuclease; 1.
CC      Pfam: PF00929; Exonuclease; 1.
CC      SMART: SM00479; EXOIII; 1.
CC      Transferrase; DNA-directed DNA polymerase; DNA replication; Hydrolase;
CC      Exonuclease; Complete proteome.
CC      SEQUENCE 229 AA; 26078 MW; CA7ADE72E63BD90 CRC64;
CC
CC      Query Match 89.7%; Score 26; DB 1; Length 229;
CC      Best Local Similarity 83.3%; Pred. No. 28;
CC      Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
CC
CC      Oy 1 REIND 6
CC      111111
CC      Db 5 REILD 10
CC
CC      RESULT 5
CC      ID ODP_RICPR STANDARD; PRT; 326 AA.
CC      AC 092DR4;
CC      DT 30-MAY-2000 (Rel. 39, Created)
CC      DT 30-MAY-2000 (Rel. 39, Last sequence update)
CC      DT 20-AUG-2001 (Rel. 40, Last annotation update)
CC      DE PYRUVATE DEHYDROGENASE E1 COMPONENT, ALPHA SUBUNIT (EC 1.2.4.1).
CC      GN PHA OR RP261.
CC      OS Rickettsia prowazekii.
CC      OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
CC      Rickettsiaceae; Rickettsia.
CC      NCBI_TaxID=782;
CC      RN 11
CC      RP SEQUENCE FROM N.A.
CC      RC STRAIN-MADRID E;
CC      RX MEDLINE-99039499; PubMed-9823893;
CC      RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
CC      Sichenitz-Ponten T., Alsmark U.C.M., Podowski R.M., Neeslund A.K.,
CC      Eriksson A.-S., Winkler H.H., Kurland C.G.;
CC      RT "The genome sequence of Rickettsia prowazekii and the origin of
CC      mitochondria."
CC      RL Nature 396:133-140(1998).
CC      -1- FUNCTION: THE PYRUVATE DEHYDROGENASE COMPLEX CATALYZES THE OVERALL
CC      CONVERSION OF PYRUVATE TO ACETYL-COA & CO(2). IT CONTAINS MULTIPLE
CC      COPIES OF THREE ENZYMAIC COMPONENTS: PYRUVATE DEHYDROGENASE (E1),
CC      DIHYDROLIPONACID ACETYLTRANSFERASE (E2) & LIPONACID DEHYDROGENASE
CC      (E3) (BY SIMILARITY)
CC      -1- CATALYTIC ACTIVITY: PYRUVATE + LIPONACID - S-ACETYL-DIHYDRO-
CC      LIPONACID + CO(2).
CC      -1- COFACTOR: THIAMINE PYROPHOSPHATE (BY SIMILARITY).
CC      -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC      -----
CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: AJ235271; CAA14723.1; -.
CC      InterPro: IPR001017; E1_ch.
CC      Pfam: PF00676; E1-dehydrog; 1.
CC      Glycylase; Oxidoreductase; Flavoprotein; Thiamine pyrophosphate;
CC      Complete proteome.
CC      SEQUENCE 326 AA; 36824 MW; BC9A6F6604B213A CRC64;
CC
CC      Query Match 89.7%; Score 26; DB 1; Length 326;
CC      Best Local Similarity 83.3%; Pred. No. 41;
CC      Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
CC
CC      Oy 1 REIND 6
CC      111111
CC      Db 277 REILD 282
CC
CC      RESULT 6
CC      ID FTSZ_BUCAI STANDARD; PRT; 384 AA.
CC      AC P57308;
CC      DT 20-AUG-2001 (Rel. 40, Created)
CC      DT 20-AUG-2001 (Rel. 40, Last sequence update)
CC      DT 20-AUG-2001 (Rel. 40, Last annotation update)
CC      DE CELL DIVISION PROTEIN FTSZ.
CC      FTSZ OR BU212.
CC      GN Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
CC      symbiotic bacterium).
CC      OS Bacteria; Proteobacteria; gamma subdivision; Buchnera.
CC      NCBI_TaxID=118099;
CC      RN 11
CC      RP SEQUENCE FROM N.A.
CC      RC STRAIN-TOKYO 1998; PubMed-10993077;
CC      RX MEDLINE-20445173; PubMed-10993077;
CC      RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
CC      RT "Genome sequence of the endocellular bacterial symbiont of aphids
CC      Buchnera sp. APS."
CC      RL Nature 407:81-86(2000).
CC      -1- FUNCTION: THIS PROTEIN IS ESSENTIAL TO THE CELL-DIVISION PROCESS.
CC      ITS SEEMS TO ASSEMBLE INTO A DYNAMIC RING ON THE INNER SURFACE OF
CC      THE CYTOPLASMIC MEMBRANE AT THE PLACE WHERE DIVISION WILL OCCUR,
CC      AND THE FORMATION OF THE RING IS THE SIGNAL FOR SEPTATION TO
CC      BEGIN. BINDS TO AND HYDROLYZES GTP (BY SIMILARITY).
CC      -1- SUBUNIT: AGGREGATE TO FORM A RING-LIKE STRUCTURE (BY SIMILARITY).
CC      -1- SUBCELLULAR LOCATION: CYTOPLASMIC; ASSEMBLE AT THE INNER SURFACE
CC      OF THE CYTOPLASMIC MEMBRANE (BY SIMILARITY).
CC      -1- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
CC      -----
CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: AP001118; BAB12929.1; -.
CC      InterPro: IPR000158; FtsZ.
CC      DR InterPro: IPR003008; Tubulin_FtsZ.
CC      Pfam: PF00091; tubulin; 1.
CC      DR PROSITE: PS01134; FTSZ_1; 1.
CC      DR PROSITE: PS01135; FTSZ_2; 1.
CC      KW Cell division, Septation; GTP-binding; Complete proteome.
CC      FT NP_BIND 103 111
CC      FT NP_BIND 103 111
CC      SEQUENCE 384 AA; 41308 MW; 477023CEBC44792 CRC64;
CC
CC      Query Match 89.7%; Score 26; DB 1; Length 384;
CC      Best Local Similarity 66.7%; Pred. No. 49;
CC      Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 RETIMD 6
11:111
DB 332 REVIMD 337

RESULT 7
NUD2_RHIME STANDARD: PRT; 404 AA.
ID NUD2_RHIME
AC P56908

DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE NADH DEHYDROGENASE I CHAIN D 2 (EC 1.6.5.3) (NADH-UBIQUINONE
OXIDOREDUCTASE CHAIN D 2).

GN NUD2.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCB1_TaxID=382;

RN (1)
RP SEQUENCE FROM N.A.

RA STRAIN=41;

RA Putnoky P., Jady B., Chellapilla K.P., Barta F., Kiss E.;

RT Rhizobium meliloti carries two sets of nuo genes.;

RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE -> NAD(+) + UBIQUINOL.

CC -1- SIMILARITY: BELONGS TO THE COMPLEX I 49 KDA SUBUNIT FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@lsb-sib.ch).

CC

CC EMBL: AJ245399; CAB51632.1;

DR InterPro: IPR001135; Complex1_49kd.

DR Pfam: PF00346; complex1_49kd; 1.

DR PROSITE: PS00535; COMPLEX1_49K; FALSE_NEG.

KM Oxidoreductase; NAD; Ubiquinone.

SO SEQUENCE 404 AA; 45675 MW; 638691610BC0A3EF CRC64;

Query Match 89.7%; Score 26; DB 1; Length 404;

Best Local Similarity 83.3%; Pred. No. 52;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RETIMD 6
11:111

DB 147 REVIMD 152

RESULT 8

Y142_NPVAC

ID Y142_NPVAC STANDARD: PRT; 477 AA.

AC P41700;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE HYPOTHECTICAL 55.4 KDA PROTEIN IN 150-151 INTERGENIC REGION.

OS Autographa californica nuclear polyhedrosis virus (AcMNPV).

OC Nucleus; dsDNA viruses, no RNA stage; Baculoviridae;

OC Nuclopolyhedrovirus.

OX NCB1_TaxID=46015;

GN (1)

RN

RP SEQUENCE FROM N.A.

RA STRAIN=C6;

RA MEDLINE=94303173; PubMed=8030224;

RA Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R.D.;

RT The complete DNA sequence of Autographa californica nuclear

RT polyhedrosis virus.;

RL Virology 202:586-605(1994).

-1- SIMILARITY: TO CORRESPONDING ORF IN OPMNPV.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@lsb-sib.ch).

CC

CC EMBL: L22858; AAA66772.1;

DR Hypothetical protein.

KM SEQUENCE 477 AA; 55417 MW; 05CEBA7AF3C91B7 CRC64;

SO

Query Match 89.7%; Score 26; DB 1; Length 477;

Best Local Similarity 83.3%; Pred. No. 62;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RETIMD 6
11:111

DB 320 REVIMD 325

RESULT 9

NIFU_BACSU

ID NIFU_BACSU STANDARD: PRT; 147 AA.

AC O32163;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE NIFU-LIKE PROTEIN.

GN NIFU

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillus/Clostridium group;

OC Bacillus/staphylococcus group; Bacillus.

OX NCB1_TaxID=1423;

GN (1)

RN

RP SEQUENCE FROM N.A.

RA STRAIN=168;

RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;

RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: MAY BE INVOLVED IN THE FORMATION OR REPAIR OF (FE-S)

CC CLUSTERS PRESENT IN IRON-SULFUR PROTEINS (POTENTIAL).

CC

CC -1- SIMILARITY: BELONGS TO THE NIFU FAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@lsb-sib.ch).

CC

CC EMBL: 299120; CAB15257.1;

DR Subtilist: BG14008; nifu.

DR InterPro: IPR002871; NIFU_N.

DR Pfam: PF01597; NIFU_N; 1.

KM Complete proteome.

SO SEQUENCE 147 AA; 16166 MW; 03F8193532B64EE2 CRC64;

Query Match 86.2%; Score 25; DB 1; Length 147;

Best Local Similarity 66.7%; Pred. No. 30;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 RETIMD 6
1:1111

DB 12 REVIMD 17

RESULT 10

PSAF_HORVU

ID PSAT HORVU STANDARD: PRT: 235 AA.
AC P13192:
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE PHOTOSYSTEM I REACTION CENTRE SUBUNIT III PRECURSOR (LIGHT-HARVESTING
DE COMPLEX I 17 KDA PROTEIN) (PSI-F).
GN PSAT.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae;
OC Trilicaceae; Hordeum.
ON NCB1_TaxID=4513;
RX NCB1_TaxID=4513;
RP SEQUENCE FROM N.A.
RA Scott M.P., Nielsen V.S., Knoetzel J., Ersen R., Moeller B.L.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE OF 78-98.
RC STRAIN-CV, PRATO;
RX MEDLINE-90033290; PubMed-2680596;
RA Anandan S., Valinstein A., Thorner J.P.;
RT *Correlation of some published amino acid sequences for photosystem I
RT polypeptides to a 17 kDa LHCI pigment-protein and to subunits III and
RT IV of the core complex.*;
RL FEBS Lett. 256:150-154(1989).
CC -1- FUNCTION: PROBABLY PARTICIPATES IN EFFICIENCY OF ELECTRON TRANSFER
CC FROM PLASTOCYANIN TO P700 (OR CYTOCHROME C553 IN ALGAE AND
CC CYNOBACTERIA). THIS PLASTOCYANIN-DOCKING PROTEIN CONTRIBUTES
CC TO THE SPECIFIC ASSOCIATION OF PLASTOCYANIN TO PS I.
CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH LUMENAL SIDE OF THE
CC THYLAKOID MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE PSAT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U08135; AAA68147.1; -;
DR PIR: S06149; S06149.
DR Mendel: 8665; HORVU:Paaf.1.
DR InterPro: IPR003666; PSI_PsaF.
DR Pfam: PF02507; PSI_PsaF.1.
KM Photosynthesis: Photosystem I; Chloroplast; Transit peptide;
KM Thylakoid; Membrane.
FT TRANSIT 1 77 CHLOROPLAST.
FT CHAIN 78 235 PHOTOSYSTEM I REACTION CENTRE SUBUNIT
FT CONFLICT 85 85 III.
FT CONFLICT 88 88 C->A (IN REF. 2).
FT CONFLICT 91 91 S->E (IN REF. 2).
FT CONFLICT 94 97 F->A (IN REF. 2).
FT CONFLICT 97 97 REKO->HEOD (IN REF. 2).
SQ SEQUENCE 235 AA; 24837 MW; 7899183C7405380A CRC64;

Query Match 86.2%; Score 25; DB 1; Length 235;
Best Local Similarity 83.3%; Pred. No. 51;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RETIID 6
DB 194 RETIID 199

RESULT 11
ID THIL_PYRO STANDARD: PRT: 309 AA.
AC 059457;
DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PROBABLE THIAMINE-MONOPHOSPHATE KINASE (EC 2.7.4.16) (THIAMINE-
DE PHOSPHATE KINASE).
GN THIL OR PH1833.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
ON NCB1_TaxID=53953;
RP SEQUENCE FROM N.A.
RC STRAIN-OT3;
RX MEDLINE-98344137; PubMed-9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Hatake Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohkubo Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT *Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.*;
RL DNA Res. 5:55-76(1998).
CC -1- CATALYTIC ACTIVITY: ATP + THIAMINE PHOSPHATE -> ADP + THIAMINE
CC DIPHOSPHATE.
CC -1- PATHWAY: THIAMINE BIOSYNTHESIS.
CC -1- SIMILARITY: BELONGS TO THE THIAMINE-MONOPHOSPHATE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP000007; BAA30953.1; -;
DR InterPro: IPR000728; ATRs-related.
DR Pfam: PF00586; ATRs.1.
KM Thiamine biosynthesis; Transferrase; Kinase; Complete proteome.
SQ SEQUENCE 309 AA; 34348 MW; 771B841B267F6511 CRC64;

Query Match 86.2%; Score 25; DB 1; Length 309;
Best Local Similarity 83.3%; Pred. No. 68;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RETIID 6
DB 122 RETIID 127

RESULT 12
ID Y376_AQUAE STANDARD: PRT: 313 AA.
AC 066696;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL PROTEIN AQ_376.
GN AQ_376.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
ON NCB1_TaxID=63363;
RP SEQUENCE FROM N.A.
RC STRAIN-VFS;
RX MEDLINE-98196666; PubMed-9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Anujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT *The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus.*;
RL Nature 392:353-358(1998).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF000687; AAC06659.1; -
 KM Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 19 41 POTENTIAL.
 FT TRANSMEM 51 68 POTENTIAL.
 FT TRANSMEM 81 103 POTENTIAL.
 SQ SEQUENCE 313 AA; 35704 MW; 3AFB57ADA036711A CRC64;

Query Match 86.2%; Score 25; DB 1; Length 313;
 Best Local Similarity 83.3%; Pred. No. 69;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 REIMD 6
 DB 266 REIMD 271

RESULT 13
 ID FVYL_HUMAN STANDARD; PRT; 578 AA.
 AC 09Y217;
 DT 20-AUG-2001 (Rel. 40; Created)
 DT 20-AUG-2001 (Rel. 40; Last sequence update)
 DT 20-AUG-2001 (Rel. 40; Last annotation update)
 DE FYVE FINGER-CONTAINING PHOSPHOINOSITIDE KINASE (EC 2.7.1.68) (1-
 DE PHOSPHATIDYLINOSITOL-4-PHOSPHATE KINASE) (PIPK) (PTDINS(4)P-5-KINASE)
 DE (P235) (FRAGMENT).
 GN PIK3VE OR KIA0981.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE:99246063; PubMed:10231032;
 RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosewa M.,
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT Prediction of the coding sequences of unidentified human genes. XIII.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.*;
 RL DNA Res. 6:63-70(1999).
 CC -1- FUNCTION: SUPPORTS THE INTRACELLULAR PIP POOL AND TO A LESSER
 CC EXTENT, THE PI 4,5-P(2) POOL. IT GENERATES PIP FROM PI AND, TO A
 CC LESSER EXTENT, PI 4,5-P(2) FROM PI 4-P. THERE ARE INDICATIONS THAT
 CC IT PHOSPHORYLATES THE D-5 RATHER THAN THE D-4 POSITION. HAS A ROLE
 CC IN ENDOSOME-RELATED MEMBRANE TRAFFICKING (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + 1-PHOSPHATIDYL-D-MYO-INOSITOL 4-
 CC MONOPHOSPHATE - ADP + 1-PHOSPHATIDYL-D-MYO-INOSITOL 4,5-
 CC BISPHOSPHATE.
 CC -1- COFACTOR: BINDS 2.2 ZINC EQUIVALENTS PER MOLECULE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PTDINS(4)P-5-KINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AB023198; BA476825.1; -
 DR EMBL: AB023198; BA476825.1; -
 DR InterPro: IPR002498; PIP5K.
 DR Pfam: PFO1504; PIP5K.1.

DR SMART: SM00330; PIPKc. 1.
 KW Transferrase; Kinase.
 FT NON_TER 1
 SQ SEQUENCE 578 AA; 64554 MW; FCB9CA1F3AFC22 CRC64;

Query Match 86.2%; Score 25; DB 1; Length 578;
 Best Local Similarity 66.7%; Pred. No. 14e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 REIMD 6
 DB 314 REIMD 319

RESULT 14
 ID EF2_HALHA STANDARD; PRT; 728 AA.
 AC P14B23;
 DT 01-APR-1990 (Rel. 14; Created)
 DT 01-APR-1990 (Rel. 14; Last sequence update)
 DT 20-AUG-2001 (Rel. 40; Last annotation update)
 DE ELONGATION FACTOR 2 (EF-2).
 DE FUSA OR FUS.
 GN Halobacterium halobium.
 OS Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;
 OC Halobacterium.
 OX NCBI_Taxid=2242;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S9.
 RX MEDLINE:90092102; PubMed:2513185;
 RT Itoh T.;
 RT *Sequence analysis of the peptide-elongation factor EF-2 gene,
 RT downstream from those of ribosomal proteins H-S12 and H-S7, from the
 RT archaeobacterial extreme halophile, Halobacterium halobium.*;
 RL Eur. J. Biochem. 186:213-219(1989).
 CC -1- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT TRANSLLOCATION
 CC OF THE NASCENT PROTEIN CHAIN FROM THE A-SITE TO THE P-SITE OF THE
 CC RIBOSOME.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
 CC EF-G/EF-2 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X17148; CA35029.1; -
 DR PIR: S07558; S07558.
 DR HSP: P07157; IAIIP.
 DR InterPro: IPR000640; EFG_C.
 DR InterPro: IPR000795; GTP-EFTU.
 DR Pfam: PF00679; EFG_C.1.
 DR Pfam: PF00009; GTP-EFTU.1.
 DR PROSITE: PS00301; EFACITOR_GTP.1.
 KW Elongation factor; GTP-binding; Protein biosynthesis.
 FT NP_BIND 28 35 GTP (BY SIMILARITY).
 FT NP_BIND 94 98 GTP (BY SIMILARITY).
 FT NP_BIND 148 151 GTP (BY SIMILARITY).
 FT MOD_RES 596 596 DIPTAMIDE (BY SIMILARITY).
 SQ SEQUENCE 728 AA; 80338 MW; EE66CB95D063D0 CRC64;

Query Match 86.2%; Score 25; DB 1; Length 728;
 Best Local Similarity 83.3%; Pred. No. 1.8e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 REIMD 6

Db 705 REIME 710

|||||

RESULT 15

MSHL_SCHPO

ID MSHL_SCHPO STANDARD; PRT: 941 AA.

AC 013921; 013700;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE MUTS PROTEIN HOMOLOG 1.

GN SPAC23C11.18C OR SPAC13F5.01C.

OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

OX NCBI_TaxID=4896;

XM [1]

RP SEQUENCE FROM N.A.

RC STRAIN=972;

RA Brown D., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;

RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: INVOLVED IN MITOCHONDRIAL DNA REPAIR (POTENTIAL).

CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTS FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: 298559; CAB1169.1; -

DR EMBL: 299091; CAB1174.1; -

DR InterPro: IPR000432; MUTS.C

DR InterPro: IPR002863; MUTS.N

DR Pfam: PF00468; MUTS.C; 1.

DR Pfam: PF01624; MUTS.N; 1.

DR ProDom: PD001263; MUTS.C; 1.

DR SMART: SM00534; MUTS.C; 1.

DR SMART: SM00533; MUTS.C; 1.

DR PROSITE: PS00486; DNA_MISMATCH_REPAIR_2; 1.

DR DNA repair; ATP-binding; DNA-binding; Mitochondrion.

KW NP_BIND 747 754 ATP (POTENTIAL).

FT NP_BIND 747 754 ATP (POTENTIAL).

SO SEQUENCE 941 AA; 106932 MW; 7A1DBF477E1140AB CRC64;

Query Match

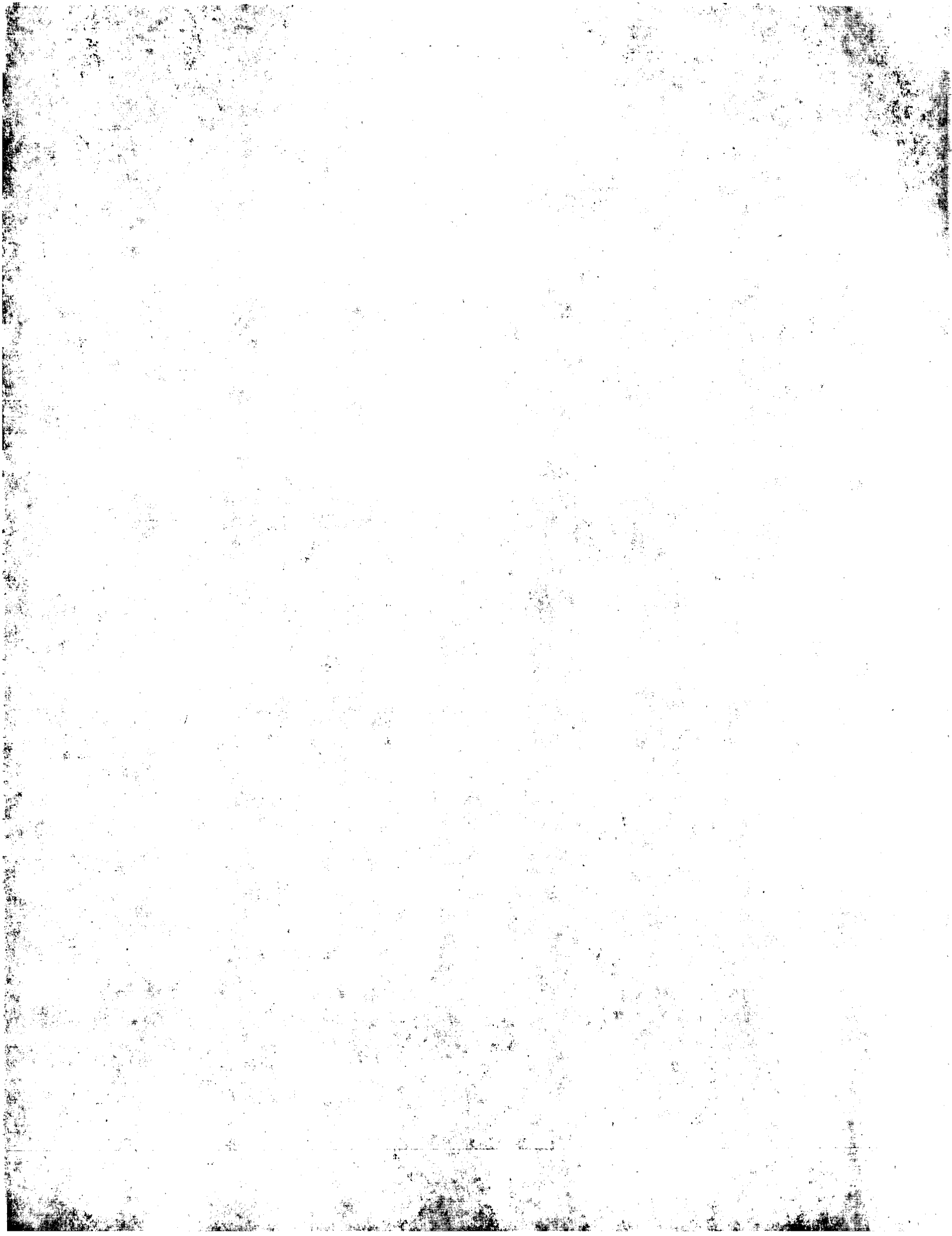
Best Local Similarity 86.2%; Score 25; DB 1; Length 941;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 REIIMD 6

DB 249 REIIVLD 254

Search completed: March 18, 2002, 09:52:27
Job time: 296 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:51:57 ; Search time 78.82 Seconds

(without alignments)
11.135 Million cell updates/sec

Title: US-09-544-665-4

Perfect score: 29

Sequence: 1 REIMD 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 segs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

SPREMBL_17:*
1: sp-archaea:*
2: sp-bacteria:*
3: sp-fungi:*
4: sp-human:*
5: sp-invertebrate:*
6: sp-mammal:*
7: sp-mhc:*
8: sp-organelle:*
9: sp-phase:*
10: sp-plant:*
11: sp-rodent:*
12: sp-virus:*
13: sp-vertebrate:*
14: sp-unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	100.0	380	1 057935	057935 methanococ
2	29	100.0	413	5 016391	016391 caenorhabdi
3	28	96.6	668	10 09NA55	09NA55 arabidopsis
4	28	96.6	817	2 09CH00	09CH00 lactococcus
5	27	93.1	243	1 059606	059606 pyrococcus
6	27	93.1	249	10 09AX28	09AX28 oryza sativ
7	27	93.1	437	1 026645	026645 methanocact
8	27	93.1	590	2 09RU01	09RU01 methanococcus
9	26	89.7	327	2 09AC56	09AC56 caulobacter
10	26	89.7	327	2 09ZCY7	09ZCY7 rickettsia
11	26	89.7	335	6 097816	097816 bos taurus
12	26	89.7	341	3 012230	012230 saccharomyc
13	26	89.7	425	12 064796	064796 autographa
14	26	89.7	445	12 091189	091189 frog adenov
15	26	89.7	476	12 092494	092494 bombyx mori
16	26	89.7	580	5 027657	027657 glardia lam
17	26	89.7	651	2 09KU86	09KU86 vibrio chol
18	26	89.7	663	6 09TU01	09TU01 bos taurus
19	26	89.7	747	3 094288	094288 schistosach

20	26	89.7	930	6 09XSW5	09XSW5 bos taurus
21	26	89.7	972	2 09RL94	09RL94 staphylococ
22	26	89.7	973	2 048348	048348 enterococcu
23	26	89.7	983	5 09V8M4	09V8M4 drosophila
24	26	89.7	1390	5 077033	077033 dicystostell
25	25	86.2	24	2 092630	092630 streptococc
26	25	86.2	24	2 092308	092308 streptococc
27	25	86.2	135	1 09HKU2	09HKU2 thermoplas
28	25	86.2	137	11 09CVS2	09CVS2 mus musculu
29	25	86.2	146	2 09K7A1	09K7A1 bacillus ba
30	25	86.2	170	5 09GR40	09GR40 caenorhabdi
31	25	86.2	175	2 085965	085965 pseudomonas
32	25	86.2	201	2 09X4W8	09X4W8 pseudomonas
33	25	86.2	236	10 09ZSU5	09ZSU5 oryza sativ
34	25	86.2	242	2 09AAZ5	09AAZ5 caulobacter
35	25	86.2	242	10 039120	039120 arabidopsis
36	25	86.2	243	2 P72510	P72510 streptococ
37	25	86.2	243	2 052232	052232 streptococ
38	25	86.2	243	2 054518	054518 streptococ
39	25	86.2	243	2 02AHN2	02AHN2 streptococ
40	25	86.2	243	2 02AHN6	02AHN6 streptococ
41	25	86.2	264	10 09C6W0	09C6W0 arabidopsi
42	25	86.2	277	10 043351	043351 arabidopsi
43	25	86.2	292	10 022851	022851 arabidopsi
44	25	86.2	311	2 099WY0	099WY0 staphylococ
45	25	86.2	313	2 066698	066698 aquifex ae

ALIGNMENTS

RESULT	ID	PRELIMINARY	PRT:	380 AA.
1	057935			
AC	057935:			
DT	01-JAN-1998 (TREMblrel. 05, Created)			
DT	01-JAN-1998 (TREMblrel. 05, Last sequence update)			
DT	01-JUN-2001 (TREMblrel. 17, Last annotation update)			
DE	PUTATIVE FORMATE HYDROGENLYASE SUBUNIT 5 (FHL SUBUNIT 5) (HYDROGENASE-3 COMPONENT E).			
CN	MT0515			
OS	Methanococcus jannaschii.			
OC	Archaea: Euryarchaeota: Methanococcales: Methanococcaceae;			
OC	Methanococcus.			
OX	NCBI_TaxID=2190;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-JAL-1 / DSM 2661 / ATCC 43067;			
RX	MEDLINE=96337999; PubMed=8688087;			
RA	Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,			
RA	Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,			
RA	Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Nguyen D.,			
RA	Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,			
RA	Uettermann T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,			
RA	Cotton M.D., Roberts K.M., Hursel M.A., Kaine B.P., Borodovsky M.,			
RA	Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;			
RT	"Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii.";			
RL	Science 273:1058-1073(1996).			
CC	-1- CONJACTOR: NICKEL (BY SIMILARITY).			
CC	-1- PATHWAY: HYDROGEN METABOLISM: FHL PATHWAY (BY SIMILARITY).			
CC	-1- SUBUNIT: FHL COMPRISSES OF A FORMATE DEHYDROGENASE, UNIDENTIFIED ELECTRON CARRIERS AND A HYDROGENASE (ISOEZYME 3). IN THIS NON-FORMATE CONSERVING PATHWAY MOLECULAR HYDROGEN AND CARBODIOXIDE FROM FORMALDEHYDE ARE RELEASED (BY SIMILARITY).			
CC	-1- SIMILARITY: BELONGS TO THE COMPLEX I 49 KDA SUBUNIT FAMILY.			
DR	EMBL: U67501; AAB9504.1; -			
DR	TIGR: M0515; -			
DR	InterPro: IPR001135; Complex1_49kd.			
DR	InterPro: IPR001501; Nifese_Hases.			
DR	Pfam: PF00346; complex1_49kd; 1.			
DR	Pfam: PF00374; Nifese_Hases; 1.			

DR PROSITE: PS00535; COMPLEX1_49K; UNKNOWN_1.
 DR PROSITE: PS00507; N1_HGNASE_L_1; UNKNOWN_1.
 KW Hypothetical protein: Oxidoreductase; NAD; Iron-sulfur; 4Fe-4S;
 SW Nickel; Complete proteome.
 SQ SEQUENCE 380 AA; 43638 MW; 3EAD16D67834A76F CRC64;

Query Match 100.0%; Score 29; DB 1; Length 380;
 Best Local Similarity 100.0%; Pred. No. 74;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 REIMD 6
 |||||
 DB 133 REIMD 138

RESULT 2
 ID 016391 PRELIMINARY; PRT: 413 AA.
 AC 016391;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE C17E7.8 PROTEIN.
 GN C17E7.8
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_Taxid=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL NZ;
 RX MEDLINE-94150718; Pubmed-7906398;
 RA Wilson R., Alscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Dublin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins J., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., Murray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken A., Roopra A., Saunders D., Shomkhen R.,
 RA Smailson N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thiermy-Mieg J., Thomas K., Vaughn M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.; et al.
 RT 2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RL elegans.
 RN Nature 368:32-38(1994).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL NZ;
 RA Bradshaw H.;
 RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL NZ;
 RA Waterston R.;
 RL Submitted (Jul-1997) to the EMBL/GenBank/DBJ databases.
 [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL NZ;
 RA Waterston R.;
 RL Submitted (Jun-1998) to the EMBL/GenBank/DBJ databases.
 CC -1 SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1 SIMILARITY: TO C4-TYPE STEROID RECEPTOR ZINC FINGER FAMILY.
 DR EMBL: AF016443; AAC4282.1;
 DR HSP: P03372; 1HCP
 DR InterPro: IPR000536; Hormone_rec_119.
 DR InterPro: IPR01628; zf-C4.
 DR Pfam: PF00104; Hormone_rec_1.
 DR Pfam: PF00105; zf-C4; 1.
 DR PRINTS: PR00047; STROIDFINGER.
 DR SMART: SM00430; HOL1; 1.
 DR SMART: SM00399; znf_C4; 1.

KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;
 KW Zinc-finger.
 SQ SEQUENCE 413 AA; 47387 MW; 91E1E6D84E702754 CRC64;

Query Match 100.0%; Score 29; DB 5; Length 413;
 Best Local Similarity 100.0%; Pred. No. 81;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 REIMD 6
 |||||
 DB 290 REIMD 295

RESULT 3
 ID 09MA55 PRELIMINARY; PRT: 668 AA.
 AC 09MA55;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE F22F7.13 PROTEIN.
 GN F22F7.13
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidops.
 OX NCBI_Taxid=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV; COLOMBIA;
 RA Lin X., Kaul S., Tom C.D., Beito M., Creasy T.H., Haas B., Wu D.,
 RA Maiti R., Rensing C.M., Koo H., Fujii C.Y., Utterback T.R.,
 RA Barnstead M.E., Bowman C.L., White O., Newman W.C., Fraser C.M.;
 RT "Arabidopsis thaliana chromosome III BAC F22F7 genomic sequence."
 DR Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC009606; AAF64540.1;
 DR InterPro: IPR000582; ACBP.
 DR InterPro: IPR001798; Kelch.
 DR Pfam: PF00887; ACBP; 1.
 DR Pfam: PF01344; Kelch; 5.
 SO SEQUENCE 668 AA; 73074 MW; F99B35A520C08ECA CRC64;

Query Match 96.6%; Score 28; DB 10; Length 668;
 Best Local Similarity 83.3%; Pred. No. 2,4e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 REIMD 6
 |||||
 DB 526 REIMD 531

RESULT 4
 ID 09CH00 PRELIMINARY; PRT: 817 AA.
 AC 09CH00;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE RIBONUCLEASE.
 GN VACB1.
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Lactococcus;
 OX NCBI_Taxid=1360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-111403;
 RA Bofeltin A., Winkler P., Manger S., Jallion O., Malarme K.,
 RA Weissbach J., Ehrlich S.D., Sorokin A.;
 RT "The complete genome sequence of the lactic acid bacterium Lactococcus
 lactis.";

RL Genome Res. 0:0-0(2001).
 DR EMBL: AE006328; AK05040.1; .
 DR InterPro: IPR001900; Ribonuclease_II.
 DR EMBL: AF003029; S1.
 DR InterPro: IPR003029; S1.
 DR Pfam: PF00575; S1.1.
 DR SMART: SM00316; S1.1.
 DR PROSITE: PS01175; RIBONUCLEASE_II; 1.
 DR Complete proteome.
 KW SEQUENCE 817 AA; 92250 MW; 44079115EF624D73 CRC64;

Query Match
 Best Local Similarity 96.6%; Score 28; DB 2; Length 817;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RETIMD 6
 DB 4 REVIMD 9

RESULT 5
 ID 059606 PRELIMINARY; PRT; 243 AA.
 AC 059606;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE HYPOTHEICAL 28.0 KDA PROTEIN PH1943.
 GN PH1943.
 OS Pyrococcus horikoshii.
 OC Archaea: Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
 OX NCBI_TaxID=53953;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OT3;
 RX MEDLINE=98344137; PubMed=9679194;
 RA Kawarabayashi Y., Sawada M., Horikawa H., Hatakeyama Y., Hino Y.,
 RA Yamamoto S., Sekine M., Baba S.-I., Koseki H., Hosoyama A., Nagai Y.,
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
 RA Funahashi T., Tanaka T., Kudo Y., Yamazaki J., Kishida N., Oguchi A.,
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
 RA Masaki Y., Shizuya H., Kikuchi H.;
 RT *Complete sequence and gene organization of the genome of a hyper-
 thermophilic archaeobacterium, Pyrococcus horikoshii OT3.*;
 RL DNA Res. 5:55-76(1998).
 DR EMBL: AP000007; BA31070.1;
 DR InterPro: IPR001993; Mtdna_carrler.
 DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_1.
 DR PROSITE: PS00037; MTD_1; UNKNOWN_1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 243 AA; 28003 MW; D643F90DA452317F CRC64;

Query Match
 Best Local Similarity 93.1%; Score 27; DB 1; Length 243;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RETIMD 6
 DB 234 RELIMD 239

RESULT 6
 ID 09AX28 PRELIMINARY; PRT; 249 AA.
 AC 09AX28;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE P0456A01.7 PROTEIN.
 GN P0456A01.7;

OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Euphorbiaceae; Oryzaeae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT Oryza sativa niponbare(GA3) genomic DNA, chromosome 1, PAC
 RT clone:P0456A01.7
 RT Submitted (Oct-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF002899; BAB21151.1;
 DR SEQUENCE 249 AA; 27391 MW; AD729DA150013E7E CRC64;

Query Match
 Best Local Similarity 93.1%; Score 27; DB 10; Length 249;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 RETIMD 6
 DB 198 REVIMD 203

RESULT 7
 ID 026645 PRELIMINARY; PRT; 437 AA.
 AC 026645;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE CONSERVED PROTEIN.
 GN MTH545.
 OS Methanobacterium thermoautotrophicum.
 OC Archaea: Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
 OX NCBI_TaxID=145262;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DELTA H;
 RX MEDLINE=98037514; PubMed=9371463;
 RA Smith D.R., Doucette-Stamm L.A., Delonghery C., Lee H.-M., Dubois J.,
 RA Hattredge T., Bashtiraden R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lum M., Pochler B., Qiu D.,
 RA Spedifora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
 RA McQuigall S., Shmer G., Goyal A., Pietrovski S., Church G.M.,
 RA Daniels C.J., Mao J.-I., Rice P., Nolling J., Reeve J.N.;
 RT *Complete genome sequence of Methanobacterium thermoautotrophicum
 deltaH: functional analysis and comparative genomics.*;
 RL J. Bacteriol 179:7135-7155(1997).
 DR EMBL: AE000837; AAB85051.1;
 KW Complete proteome.
 SQ SEQUENCE 437 AA; 48978 MW; 04D3591401BA199A CRC64;

Query Match
 Best Local Similarity 93.1%; Score 27; DB 1; Length 437;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RETIMD 6
 DB 63 RELIMD 68

RESULT 8
 ID 09RU01 PRELIMINARY; PRT; 590 AA.
 AC 09RU01;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

```

DE GLUCOSE-6-PHOSPHATE 1-DEHYDROGENASE (EC 1.1.1.49) (66PD).
GN DRI596.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moftat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Uitterback T., Zaleski C.,
RA Makarewicz K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioreistant bacterium Deinococcus
RT radiodurans R1."
RL Science 286:1571-1577(1999).
CC -1- CATALYTIC ACTIVITY: D-GLUCOSE 6-PHOSPHATE + NADP(+) = D-GLUCONO-
CC DELTA-LACTONE 6-PHOSPHATE + NADPH.
CC -1- PATHWAY: FIRST STEP IN PENTOSE PHOSPHATE PATHWAY.
CC -1- SIMILARITY: TO THE GLUCOSE-6-PHOSPHATE DEHYDROGENASE FAMILY.
DR EMBL; AE002003; AAF1158.1;
DR HSSP; P11411; 1DPG.
DR TIGR; DRI596;
DR InterPro; IPR001282; G6PD.
DR Pfam; PF00479; G6PD. 1.
DR PRODOM; PD001129; G6PD. 1.
DR PROSITE; PS00069; G6P_DEHYDROGENASE; 1.
KW Complete proteome: Glucose metabolism; NADP: Oxidoreductase.
SQ SEQUENCE 590 AA; 65600 MW; DA42AEBBD277722 CRC64;

Query Match 93.1%; Score 27; DB 2; Length 590;
Best Local Similarity 66.7%; Pred. No. 3.6e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 REIMD 6
Db 500 REVMD 505

RESULT 9
ID 09AC56 PRELIMINARY; PRT; 237 AA.
AC 09AC56;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE DNA POLYMERASE III, EPSILON SUBUNIT.
GN CC0005.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group.
OC Caulobacter.
OX NCBI_TaxID=69394;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Uitterback T., Tran K., Wolf A., Vamathevan J., Emswiler M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus."
RT Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AE005675; AKK21993.1;
DR TIGR; CC0005;
KW Complete proteome.
SQ SEQUENCE 237 AA; 26459 MW; 96241B905E74CD5D CRC64;

```

```

Query Match 89.7%; Score 26; DB 2; Length 237;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 REIMD 6
Db 3 REIMD 8

RESULT 10
ID 09ZCY7 PRELIMINARY; PRT; 327 AA.
AC 09ZCY7;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE HYPOTHETICAL 37.5 KDA PROTEIN.
GN RP564.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales.
OC Rickettsiaceae; Rickettsiidae; Rickettsia.
OX NCBI_TaxID=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MADRID E;
RX MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Sierichitz-Ponten T., Almark U.C.M., Podowski R.M., Naeslund A.K.,
RA Erlanson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria."
RL Nature 396:133-140(1998).
DR EMBL; AJ235272; CAH15012.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 327 AA; 37453 MW; 1F1672BF492FEFEF CRC64;

Query Match 89.7%; Score 26; DB 2; Length 327;
Best Local Similarity 83.3%; Pred. No. 3.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 REIMD 6
Db 153 ROLIMD 158

RESULT 11
ID 097916 PRELIMINARY; PRT; 335 AA.
AC 097916;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE REVERSE TRANSCRIPTASE-LIKE.
GN RTF.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Mierinckx A., Mercer D., Oulmouden A., Petit J.M., Julien R.;
RT "Complete genomic organization of bovine fub gene reveals that human
RT 5'-untranslated exons of FUT3 and FUT6 have recently emerged from
RT ancestral intronic sequences."
RT Submitted (FEB-1999) to the EMBL/GenBank/DDP databases.
CC -1- SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE
CC TRANSCRIPTASE).
DR EMBL; AJ132772; CA10770.1;
DR InterPro; IPR000477; RVTse.

```

DR Pfam: PF00078; rvc: 1.
 KW RNA-directed DNA polymerase.
 SQ SEQUENCE 335 AA; 38403 MW; C7A8857EF80F6D23 CRC64;

Query Match 89.7%; Score 26; DB 6; Length 335;
 Best Local Similarity 66.7%; Pred. No. 3.4e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 RETIIMD 6
 DB 310 RELVMD 315

RESULT 12
 ID 012230 PRELIMINARY; PRT; 341 AA.
 AC 012230:
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DE 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
 DE HYPOTHETICAL 38.1 KDA PROTEIN YPL004C.
 GN YPL004C OR YPLA13C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AB972;
 RA Badcock K., Bowman S., Churcher C.M., Pearson D., Rajandream M.A.,
 RA Walsh S.V., Barrell B.G.;
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AB972;
 RA Badcock K., Churcher C.;
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AB972;
 RA Walsh S.V., Barrell B., Rajandream M.A.;
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 MEDLINE=89057455; PubMed=3143101;
 RA O'Hara P.J., Horowitz H., Eichinger G., Young E.T.;
 RT "The yeast ADR6 gene encodes homopolymeric amino acid sequences and a
 RT potential metal-binding domain.";
 RL Nucleic Acids Res. 16:10153-10169(1988).
 RN [5]
 RP SEQUENCE FROM N.A.
 MEDLINE=91092236; PubMed=2265610;
 RA Gertling S.L., Spencer F., Hlester P.;
 RT "The CHL 1 (CTF 1) gene product of Saccharomyces cerevisiae is
 RT important for chromosome transmission and normal cell cycle
 RT progression in G2/M.";
 RL EMBO J. 9:4347-4358(1990).
 RN [6]
 RP SEQUENCE FROM N.A.
 RA Hall J., Ahmed A., Bussey H., Fortin N., Friesen J.D., Storms R.K.,
 RA Vo D.H., Wang Y., Minnett E.,
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z71253; CA95037.1;
 DR EMBL; Z48483; CA88382.1;
 DR EMBL; U33335; AAB6101.1;
 DR SGD; S0005925; YPL004C.
 KW Hypothetical protein.
 SO SEQUENCE 341 AA; 38071 MW; 6FCCE022ECC9778C CRC64;

Query Match 89.7%; Score 26; DB 3; Length 341;
 Best Local Similarity 83.3%; Pred. No. 3.5e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RETIIMD 6
 DB 248 RETIIMD 253

RESULT 13
 ID 064796 PRELIMINARY; PRT; 425 AA.
 AC 064796:
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-AUG-1998 (TRENBLrel. 07, Last annotation update)
 DE EARLY 49 KDA PROTEIN.
 OS Autographa californica nuclear polyhedrosis virus (ACNPV).
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_TaxID=46015;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=E2;
 RA Ramamurthy P., Braunagel S.C., Summers M.D.;
 RL Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U04051; AAA03635.1;
 SQ SEQUENCE 425 AA; 49374 MW; 7572BD3EBC6B2A52 CRC64;

Query Match 89.7%; Score 26; DB 12; Length 425;
 Best Local Similarity 83.3%; Pred. No. 4.4e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RETIIMD 6
 DB 268 RETIIMD 273

RESULT 14
 ID 091H9 PRELIMINARY; PRT; 445 AA.
 AC 091H9:
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
 DE I11 PROTEIN.
 OS Frog adenovirus 1.
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Sladenovirus.
 OX NCBI_TaxID=114102;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Davison A.J., Wright K.M., Harrach B.;
 RT "Phylogenetic position of an amphibian adenovirus.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF224336; AAF6928.1;
 DR InterPro; IPR002605; Adeno_Penton_B.
 DR Pfam; PF01686; Adeno_Penton_B; 1.
 SO SEQUENCE 445 AA; 50029 MW; E5061EFB041CB0CD CRC64;

Query Match 89.7%; Score 26; DB 12; Length 445;
 Best Local Similarity 83.3%; Pred. No. 4.7e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RETIIMD 6
 DB 71 RETIIMD 76

RESULT 15
 ID 092494 PRELIMINARY; PRT; 476 AA.
 AC 092494:
 DT 01-NOV-1998 (TRENBLrel. 08, Created)

DT 01-NOV-1998 (TrEMBLrel. 08, last sequence update)
 DT 01-NOV-1998 (TrEMBLrel. 08, last annotation update)
 DE ACMPV ORF142.
 OS Bombyx mori nuclear polyhedrosis virus (BmNPV).
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_TaxID-10458;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-T3;
 RX MEDLINE-97329351; PubMed-9185864;
 RA Kamita S.G., Maeda S.;
 RT "Sequencing of the putative DNA helicase-encoding gene of the Bombyx
 mori nuclear polyhedrosis virus and fine-mapping of a region involved
 in host range expansion.";
 RL Gene 190:173-179(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-T3;
 RA Gomi S., Majima K., Maeda S.;
 RT "Sequence analysis of the genome of Bombyx mori
 nucleopolyhedrovirus.";
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U33180; AAC63808.1; -;
 SQ SEQUENCE 476 AA; 55375 MW; AC0E601022609068 CRC64;

Query Match 89.7%; Score 26; DB 12; Length 476;
 Best Local Similarity 83.3%; Pred. No. 5e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 REITMD 6
 Db 320 REITLD 325

Search completed: March 18, 2002, 09:51:59
 Job time: 288 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:48:56 ; Search time 81.32 Seconds

(without alignments)
5.465 Million cell updates/sec

Title: US-09-544-665-5

Perfect score: 28

Sequence: 1 EEIMR 6

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: /SIDS8/gcgdata/geneseq/AA1980.DAT:*
2: /SIDS8/gcgdata/geneseq/AA1981.DAT:*
3: /SIDS8/gcgdata/geneseq/AA1982.DAT:*
4: /SIDS8/gcgdata/geneseq/AA1983.DAT:*
5: /SIDS8/gcgdata/geneseq/AA1984.DAT:*
6: /SIDS8/gcgdata/geneseq/AA1985.DAT:*
7: /SIDS8/gcgdata/geneseq/AA1986.DAT:*
8: /SIDS8/gcgdata/geneseq/AA1987.DAT:*
9: /SIDS8/gcgdata/geneseq/AA1988.DAT:*
10: /SIDS8/gcgdata/geneseq/AA1989.DAT:*
11: /SIDS8/gcgdata/geneseq/AA1990.DAT:*
12: /SIDS8/gcgdata/geneseq/AA1991.DAT:*
13: /SIDS8/gcgdata/geneseq/AA1992.DAT:*
14: /SIDS8/gcgdata/geneseq/AA1993.DAT:*
15: /SIDS8/gcgdata/geneseq/AA1994.DAT:*
16: /SIDS8/gcgdata/geneseq/AA1995.DAT:*
17: /SIDS8/gcgdata/geneseq/AA1996.DAT:*
18: /SIDS8/gcgdata/geneseq/AA1997.DAT:*
19: /SIDS8/gcgdata/geneseq/AA1998.DAT:*
20: /SIDS8/gcgdata/geneseq/AA1999.DAT:*
21: /SIDS8/gcgdata/geneseq/AA2000.DAT:*
22: /SIDS8/gcgdata/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	100.0	6	20	AAV01814
2	27	96.4	401	20	AAE95362
3	27	96.4	1757	20	AAH84351
4	27	96.4	1757	20	AAH84351
5	27	96.4	2763	18	AAV10344
6	25	89.3	246	22	AAE98338
7	25	89.3	859	16	AAH82886
8	25	89.3	859	18	AAW31227
9	25	89.3	1132	21	AAE47090
10	25	89.3	1176	21	AAE47089
11	25	89.3	1238	21	AAE47088

12	24	85.7	265	20	AAH93359
13	24	85.7	306	16	AAE9825
14	24	85.7	306	19	AAE60180
15	24	85.7	314	21	AAE81702
16	24	85.7	329	21	AAE81704
17	24	85.7	378	19	AAH98595
18	24	85.7	811	20	AAE29146
19	24	85.7	846	11	AAH8406
20	24	85.7	846	11	AAE6353
21	24	85.7	853	19	AAH43077
22	24	85.7	859	19	AAH43080
23	24	85.7	860	21	AAE69348
24	24	85.7	961	20	AAE34577
25	24	85.7	973	20	AAE34434
26	24	85.7	1698	22	AAE62448
27	24	82.1	6	20	AAV01812
28	23	82.1	19	13	AAE26000
29	23	82.1	33	22	AAE42738
30	23	82.1	33	22	AAE82200
31	23	82.1	47	20	AAE12851
32	23	82.1	137	21	AAE50790
33	23	82.1	138	20	AAE12378
34	23	82.1	152	21	AAE01421
35	23	82.1	187	20	AAE48261
36	23	82.1	198	20	AAE20033
37	23	82.1	217	18	AAE18063
38	23	82.1	220	21	AAE32041
39	23	82.1	223	20	AAE20032
40	23	82.1	236	21	AAE32040
41	23	82.1	240	22	AAE01058
42	23	82.1	273	21	AAE33339
43	23	82.1	281	17	AAE47105
44	23	82.1	281	17	AAE08083
45	23	82.1	282	17	AAE08084

ALIGNMENTS

RESULT 1	
ID	AAV01814 standard; peptide: 6 AA.
XX	
AC	AAV01814:
XX	
DT	29-JUN-1999 (first entry)
XX	
DE	Peptide used to regulate scupa.
XX	
KW	Abnormal cell migration; PAI-1-dependent cell adhesion; scupa clearance;
KW	pathological cell migration; angiogenesis; organogenesis; ovulation;
KW	inflammation; cancer; tumor cell invasion; metastasis; atherosclerosis.
OS	Synthetic.
XX	
PN	W09920295-A1.
XX	
PD	29-APR-1999.
XX	
PF	15-OCT-1998; 98MO-US21800.
XX	
PR	17-OCT-1997; 97US-0062274.
XX	
PA	(UNP-) UNIV PENNSYLVANIA.
XX	
PI	Clines D, Hlgazi AA;
XX	
DR	WPI: 1999-288168/24.
XX	
PT	Peptide composition
XX	
PS	Example 1; Page 31; 63pp; English.
XX	

Canine interleukin
transposon th3401
B. thuringiensis t
Streptococcus pneu
Streptococcus pneu
H. pylori GHP0.117
Amino acid sequenc
Sequence deduced f
HIV-1 non-subtype
HIV-1 gp120 protei
HIV-1 gp120 protei
HIV-1 non-subtype
Porphyromonas ging
Porphyromonas ging
S. cerevisiae YDRI
Peptide used to re
N-terminal t-PA in
Human alpha-1-anti
Thrombin inhibitor
Human 5' EST secre
Retroltransposon co
Human 5' EST secre
Human secreted pro
Human prostate can
B. burgdorferi ant
Growth factor rece
Arabidopsis thalia
Arabidopsis thalia
B. burgdorferi ant
Arabidopsis thalia
CEE 61 protein seq
Arabidopsis thalia
Arabidopsis thalia
Mouse ST-S1, used
Human ST-S1, used

CC Peptides AAY01812-14 exemplify the peptide of the invention, which is
CC used in a peptide composition to affect a biological process
CC characterized by abnormal cell migration through a physiological barrier,
CC to inhibit PAI-1-dependent adhesion of a cell to a tissue, to promote
CC clearance of scupa from the surface of a mammalian cell, and to impede
CC pathological migration of a cell in a mammal. The composition can be used
CC to affect a biological process characterized by abnormal cell migration
CC through a physiological barrier, where the process is selected from cell
CC angiogenesis, organogenesis, ovulation, inflammation, cancer, tumor cell
CC invasion and metastasis, and atherosclerosis.
XX
SQ Sequence 6 AA:

Query Match 100.0%; Score 28; DB 20; Length 6;
Best Local Similarity 100.0%; Pred. No. 4.3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6
DB 1 eelimr 6

RESULT 2

AAB95262
ID AAB95262 standard; Protein; 401 AA.

XX AAB95262;

XX 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:17446.

XX Human; primer: detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-018776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI: 2001-318749/34.

XX Claim 8: SEQ ID 17446; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602
XX full-length cDNAs defined in the specification. Where a primer set
XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary
XX to the complementary strand of a polynucleotide which comprises one of
XX the 5602 nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the

CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialized methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95993 represent human amino acid sequences; and AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 401 AA:

Query Match 96.4%; Score 27; DB 22; Length 401;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6
DB 306 eelimr 311

RESULT 3

AAM84351
ID AAM84351 standard; Protein; 1757 AA.

XX AAM84351;

XX 25-MAR-1999 (first entry)

XX Murine ubiquitin-protein ligase Ubr1.

XX Ubiquitin-protein ligase; Ubr1; mouse; ubiquitinylation; degradation;
XX N-end rule pathway; stress-related muscle wasting; inhibitor; screen.

XX Mus sp.

XX U55861312-A.

XX 19-JAN-1999.

XX 02-DEC-1997; 97US-0982956.

XX 02-DEC-1997; 97US-0982956.

XX (CALY) CALIFORNIA INST OF TECHNOLOGY.

XX Kwon YT, Varshavsky A;

XX WPI: 1999-130395/11.

XX N-PSDB: AAV9308.

XX Mouse and human Ubr1 cDNA - useful for producing recombinant Ubr1
XX polypeptides

XX Disclosure: Columns 15-28; 18pp; English.

XX The present sequence represents a ubiquitin-protein ligase called Ubr1.
XX The Ubr1 enzymes are involved in protein ubiquitinylation and
XX ultimate degradation through the N-end rule pathway and have been
XX linked to stress-related muscle wasting. Recombinant Ubr1 polypeptides
XX can be used to screen for inhibitors of muscle wasting when this is
XX associated with the N-end rule pathway.
XX
SQ Sequence 1757 AA:

Query Match 96.4%; Score 27; DB 20; Length 1757;
Best Local Similarity 83.3%; Pred. No. 6.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EETIMR 6
11:111
Db 765 eevimr 770

RESULT 4

AAB31162
ID AAB31162 standard; Protein; 1757 AA.

AC AAB31162;

DT 02-APR-2001 (first entry)

XX Amino acid sequence of Mouse Ubr1 protein.

XX Ubr1: E3-type protein; ubiquitin system; ubiquitin-protein ligase;

KW N-end rule pathway; intracellular pathogen; Lysteria monocytogenes;

KW Yersinia enterocolitica; muscle wasting; infection.

OS Mus sp.

PN US6159732-A.

PD 12-DEC-2000.

PF 11-JAN-1999; 99US-0228317.

PR 02-DEC-1997; 97US-0982956.

PA (CALY) CALIFORNIA INST OF TECHNOLOGY.

PI Kwon YT, Varshavsky A;

DR WPI: 2001-090278/10.

DR N-PSDB: AAC86933.

XX Inhibiting the N-end rule pathway in mammalian cells for treating

PT infections and various diseases associated with muscle tissue wasting,

PT by inhibiting the expression of Ubr1 gene

XX Example: Column 15-28; 18pp; English.

XX The present sequence represents a murine Ubr1 enzyme. Ubr1 is an E3-type

CC protein of the ubiquitin system. Specifically, it is a ubiquitin-protein

CC ligase. The enzyme is specific for destabilizing residues exposed at

CC the N-terminus of protein substrates. Inhibition of the expression of

CC Ubr1 gene in a cell results in inhibition of the N-end rule pathway.

CC The method is used for treatment of mammalian cells infected with an

CC intracellular pathogen, e.g. Lysteria monocytogenes or Yersinia

CC enterocolitica. Inhibition of N-end rule pathway is also useful for

CC treating various diseases associated with wasting of muscle tissue and

XX Sequence 1757 AA;

Db 765 eevimr 770

RESULT 5

AAAM10344
ID AAM10344 standard; Protein; 2763 AA.

XX AAM10344;

DT 15-APR-1997 (first entry)

XX Maize dwarf mosaic virus polyprotein.

DE MDMV-B: viral resistance; disease resistance; transgenic plant;

KW monocot; P3 proteinase; N1a proteinase; N1b replicase;

KW RNA-dependent RNA polymerase; coat protein; zea mays; sorghum;

XX sugarcane; Saccharum officinale.

XX Maize dwarf mosaic virus strain B.

OS Key

XX Location/Qualifiers

XX 1..377

XX /label=HC-Pro

XX /note="portion of the helper component-P2

XX proteinase"

XX Domain

XX 263..336

XX /note="MDMV-B HC-Pro domain"

XX Protein

XX 378..791

XX /label=P3-proteinase

XX /note="claimed polypeptide (claim 20)"

XX Protein

XX 792..1430

XX /label=Cylindrical_inclusion_protein

XX /note="claimed polypeptide (claim 20)"

XX Domain

XX 880..1010

XX /note="conserved helicase domain"

XX Protein

XX 1431..1483

XX /label=X2

XX /note="6 kda protein"

XX Protein

XX 1484..1914

XX /label=N1a-Proteinase

XX /note="claimed polypeptide (claim 20)"

XX Protein

XX 1915..2435

XX /label=N1b-replicase

XX /note="claimed polypeptide (claim 20)"

XX Region

XX 2266..2268

XX /note="conserved motif characteristic of

XX RNA-dependent replicases"

XX Protein

XX 2436..2763

XX /label=Coat_protein

XX Misc-difference 144

XX /note="unidentified amino acid"

XX Misc-difference 704

XX /note="unidentified amino acid"

XX Misc-difference 712

XX /note="unidentified amino acid"

XX Misc-difference 829

XX /note="unidentified amino acid"

XX Misc-difference 834

XX /note="unidentified amino acid"

XX Misc-difference 843

XX /note="unidentified amino acid"

XX Misc-difference 847

XX /note="unidentified amino acid"

XX Misc-difference 852

XX /note="unidentified amino acid"

XX Misc-difference 1346

XX /note="unidentified amino acid"

XX Misc-difference 1362

XX /note="unidentified amino acid"

XX Misc-difference 2077

XX /note="unidentified amino acid"

XX Misc-difference 2385

XX /note="unidentified amino acid"

XX W09702352-A1.

XX 23-JAN-1997.

XX 20-JUN-1996; 96WO-EP02673.

XX 30-JUN-1995; 95US-0496944.

XX (CIBA) CIBA GEIGY AG.

XX Dietz JM, Law MD.
XX
XX WPI: 1997-108965/10.
DR N-PSDB: AAT747073.
XX
XX Chimeric gene for imparting viral resistance to plants - contains
PT sequence modified to express non-translatable mRNA, or non-coat
PT viral protein
XX
XX Claim 20: Page 31-44: 64pp: English.
XX
XX The sequence of the polyprotein encoded by the polycistronic mRNA
CC (AAT747073) of maize dwarf mosaic virus strain B (MDMV-B) is given in
CC AAT10344. New chimeric genes comprise a monocotyledonous plant
CC promoter linked to a modified nucleic acid sequence derived from
CC the MDMV-B genome. The modification is such that mRNA is translated
CC to a truncated protein (pref. smaller than 200 amino acids), no
CC translation of mRNA occurs or the transcribed mRNA lacks the
CC translation initiation codon or includes a premature stop codon.
CC Expression of the chimeric gene inhibits infection of plants (pref.
CC sorghum, sugarcane, esp. maize) by MDMV. The transgenic plants
CC display an inheritable resistance trait.
CC
SQ Sequence 2763 AA:

Query Match 96.4%; Score 27; DB 18; Length 2763;
Best Local Similarity 83.3%; Pred. No. 1.1e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6
Db 2064 eelivr 2069

RESULT 6
AAC98338
ID AAC98338 standard; Protein: 246 AA.
XX
AC AAC98338:
XX
DT 21-SEP-2001 (first entry)
XX
DE Escherichia coli protein sequence SEQ ID NO:386.
XX
XX Escherichia coli; identification; proliferation; microorganism;
KW antimicrobial; antibacterial; antibiotic; gene therapy; diagnosis;
KW bacterial growth inhibition.
XX
OS Escherichia coli.
XX
PN MO200148209-A2.
XX
PD 05-JUL-2001.
XX
PF 19-DEC-2000; 2000MO-US34419.
XX
PR 23-DEC-1999; 99US-0173005.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Forsyth RA, Ohlsen KL, Zyskind JM.
XX
XX WPI: 2001-457376/49.
DR N-PSDB: AAH81394.
XX
PT Novel nucleic acids encoding proteins required for Escherichia coli
PT proliferation, useful for screening for antimicrobial agents -
XX Claim 19; Page 507-508; 596pp: English.
XX
XX The present invention describes a purified or isolated nucleic acid

CC sequence (I) consisting essentially of one of the 93 nucleotide sequences
CC given in AAH81202 to AAH81294, where expression of the nucleic acid in a
CC microorganism is capable of inhibiting proliferation of a microorganism.
CC (I) have antibacterial and antibiotic activities, and can be used in
CC gene therapy. Expression of (I) in a microorganism inhibits proliferation
CC of the microorganism, and the manufactured antibiotic is useful for
CC reducing the activity or level of a gene product required for
CC proliferation of a microorganism in a subject, specifically humans. The
CC nucleic acids that inhibit bacterial growth or proliferation can be used
CC as antisense therapeutics for killing bacteria. In addition to
CC therapeutic applications, the nucleic acid sequences complementary to
CC sequences required for proliferation can be used as diagnostic tools.
CC For example, nucleic acid probes complementary to proliferation-required
CC sequences that are specific for particular species of microorganisms can
CC be used as probes to identify particular microorganism species in
CC clinical specimens. AAH81295 to AAH81487 encode the Escherichia coli
CC proteins given in AAC98239 to AAC98431, and AAH81488 to AAH81491
CC represent oligonucleotides, which are used in the exemplification of the
CC present invention.
XX
SQ Sequence 246 AA:

Query Match 89.3%; Score 25; DB 27; Length 246;
Best Local Similarity 83.3%; Pred. No. 2.0e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6
Db 173 qelivr 178

RESULT 7
AAR82886
ID AAR82886 standard; Protein: 859 AA.
XX
AC AAR82886:
XX
DT 10-FEB-1996 (first entry)
XX
DE Human leucine zipper protein-kinase.
XX
KW Leucine zipper protein-kinase; enzyme; zpk; CNS; diagnosis;
KW central nervous system.
XX
OS Homo sapiens.
XX
XX MO9523849-A1.
XX
PD 08-SEP-1995.
XX
PF 28-FEB-1995; 95MO-US02792.
XX
PR 01-MAR-1994; 94US-0205018.
XX
PA (CHIL-) CHILRENS HOSPITAL PHILADELPHIA.
XX
PI Pleasure DE, Reddy U;
XX
XX WPI: 1995-320565/41.
DR N-PSDB: AAT01031.
XX
XX DNA encoding novel leucine zipper protein kinase - also probe(s) and
PT polypeptide(s) prepared using the DNA, useful for inhibiting
PT hyper-proliferation of CNS cells
XX
XX Claim 4; Page 23-27; 40pp: English.
XX
XX This novel protein may be used to treat tumors of the CNS in a
CC mammal by inhibiting the overexpression of the leucine zipper
CC protein-kinase in vivo, or by interfering with a vital signal
CC in a chain of signals leading to tumorigenicity.
XX

SQ Sequence 859 AA:

Query Match 89.3%; Score 25; DB 16; Length 859;
 Best Local Similarity 66.7%; Pred. No. 9.4e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6
 11::11
 DB 409 eelvmr 414

RESULT 8

AAW31227

ID AAW31227 standard; Protein: 859 AA.

AC AAW31227;

DT 17-FEB-1998 (first entry)

DE Human leucine-zipper protein kinase.

KW Leucine zipper protein kinase; zpk; human; metastasis; tumour;

KW serine/threonine protein kinase; non-receptor type kinase;

KW cell hyperproliferation.

OS Homo sapiens.

FH Key

Location/Qualifiers
1..859

FT Protein

/label- zpk
/note- "leucine zipper protein kinase"

FT Domain

231..243
/label- protein_kinase_domain
/note- "As stated in specification"

FT Misc-difference 234..235

/note- "Mentioned in specification"

FT Misc-difference 236..237

/note- "Mentioned in specification"

FT Misc-difference 240

/note- "Mentioned in specification"

FT Misc-difference 251

/note- "Mentioned in specification"

FT Misc-difference 254..256

/note- "Mentioned in specification"

FT Misc-difference 278..280

/note- "Mentioned in specification"

FT Misc-difference 292

/note- "Mentioned in specification"

FT Misc-difference 294..295

/note- "Mentioned in specification"

FT Misc-difference 297

/note- "Mentioned in specification"

FT Region

415..418
/note- "Mentioned in specification"

FT Region

442..468
/label- Leucine zipper motif
/note- "Putative endoplasmic reticulum targeting sequence as given in the specification"

FT Misc-difference 443

/note- "As stated in specification"

FT Misc-difference 443

/note- "Mentioned in specification"

FT Misc-difference 450

/note- "Mentioned in specification"

FT Misc-difference 457

/note- "Mentioned in specification"

FT Misc-difference 464

/note- "Mentioned in specification"

FT Region

537..544
/label- ATP binding site
/note- "As stated in specification"

XX US5676945-A.

PD 14-OCT-1997.

XX 01-MAR-1994; 94US-0205018.

XX 28-FEB-1995; 95US-0395580.

XX 01-MAR-1994; 94US-0205018.

XX (CHIL-) CHILDRENS HOSPITAL PHILADELPHIA.

XX Pleasure D. Reddy U;

XX WPI; 1997-511822/47.

XX N-PSDB; AAT89349.

XX Human leucine-zipper protein kinase - useful for treating tumours of the central nervous system

XX Claim 3; Fig 1; 19pp; English.

CC This sequence represents a novel human leucine-zipper protein kinase isolated from brain tissue. The specification states that the protein contains an ATP-binding site at position 537-544 (consensus

CC sequence Gly-Xaa-Gly-Xaa-Gly), a protein kinase domain at position 231-243 and a putative endoplasmic reticulum (ER) targeting sequence at position 413-418 (consensus sequence RDEL). This protein is most similar

CC to members of serine/threonine protein kinases and is believed to be a "non-receptor type kinase" based on its lack of a transmembrane domain.

CC probes to this protein could be used for diagnostic or research purposes to detect or quantitate the expression of leucine-zipper protein kinase.

CC Overexpression of leucine zipper protein kinase can result in hyperproliferation of cells and metastasis. The application of exogenous

CC leucine-zipper protein kinase may interfere with specific protein-protein or protein-nucleic acid interactions involved in hyperproliferation. This

CC may be used to treat animals suffering from tumours of the central nervous system by inhibiting the overexpression of leucine-zipper protein

CC kinase in vivo or by interfering with a vital signal in the chain of signals leading to tumorigenicity.

CC signals leading to tumorigenicity.

SQ Sequence 859 AA:

Query Match 89.3%; Score 25; DB 18; Length 859;
 Best Local Similarity 66.7%; Pred. No. 9.4e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6
 11::11
 DB 409 eelvmr 414

RESULT 9

AAG47090

ID AAG47090 standard; Protein: 1132 AA.

AC AAG47090;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 59314.

KW Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.

KW Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123588.
PR 23-MAR-1999; 99US-0125788.
PR 23-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128242.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134286.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137552.
PR 04-JUN-1999; 99US-0137582.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140635.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140921.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0143277.
PR 13-JUL-1999; 99US-0143342.

PR 14-JUL-1999; 99US-0144624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147139.
PR 05-AUG-1999; 99US-0147266.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147418.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 26-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.

PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0158294.
PR 13-OCT-1999; 99US-0158293.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 89.3% Score 25; DB 21; Length 1132;
Best Local Similarity 83.3% Pred. No. 1,2e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6
Db 211 eelink 216

RESULT 10

ID AAG47089 standard; Protein; 1176 AA.

AC AAG47089;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 59313.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

OS Arabidopsis thaliana.

PN EPI033405-A2.

XX 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125786.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132464.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135114.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139859.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142380.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.

```

PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0150085.
PR 22-JUL-1999; 99US-0150087.
PR 22-JUL-1999; 99US-0150889.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145115.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0148171.
PR 10-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148664.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149733.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0156456.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157133.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158233.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.

```

```

PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      89.3%; Score 25; DB 21; Length 1176;
Best Local Similarity 83.3%; Pred. No. 1.3e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EELIMR 6
Db 255 eelink 260

RESULT 11
AA647088
ID AA647088 standard; Protein; 1238 AA.
XX
AC AA647088;
XX
DT 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 59312.
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 59312.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.

```

PR	30-APR-1999	9905-0132048
PR	30-APR-1999	9905-0132047
PR	04-MAY-1999	9905-0132484
PR	05-MAY-1999	9905-0132485
PR	06-MAY-1999	9905-0132486
PR	06-MAY-1999	9905-0132487
PR	07-MAY-1999	9905-0132488
PR	11-MAY-1999	9905-0132489
PR	14-MAY-1999	9905-0134216
PR	14-MAY-1999	9905-0134221
PR	14-MAY-1999	9905-0134370
PR	18-MAY-1999	9905-0134768
PR	19-MAY-1999	9905-0134914
PR	20-MAY-1999	9905-0135124
PR	21-MAY-1999	9905-0135333
PR	24-MAY-1999	9905-0135639
PR	25-MAY-1999	9905-0136021
PR	27-MAY-1999	9905-0136782
PR	01-JUN-1999	9905-0137222
PR	03-JUN-1999	9905-0137502
PR	04-JUN-1999	9905-0137507
PR	07-JUN-1999	9905-0137724
PR	08-JUN-1999	9905-0138094
PR	10-JUN-1999	9905-0138504
PR	14-JUN-1999	9905-0138847
PR	16-JUN-1999	9905-0139119
PR	16-JUN-1999	9905-0139452
PR	17-JUN-1999	9905-0139453
PR	18-JUN-1999	9905-0139492
PR	18-JUN-1999	9905-0139454
PR	18-JUN-1999	9905-0139455
PR	18-JUN-1999	9905-0139456
PR	18-JUN-1999	9905-0139457
PR	18-JUN-1999	9905-0139458
PR	18-JUN-1999	9905-0139459
PR	18-JUN-1999	9905-0139460
PR	18-JUN-1999	9905-0139461
PR	18-JUN-1999	9905-0139462
PR	18-JUN-1999	9905-0139463
PR	18-JUN-1999	9905-0139750
PR	18-JUN-1999	9905-0139763
PR	21-JUN-1999	9905-0139817
PR	22-JUN-1999	9905-0139819
PR	23-JUN-1999	9905-0140353
PR	23-JUN-1999	9905-0140354
PR	24-JUN-1999	9905-0140655
PR	28-JUN-1999	9905-0140823
PR	29-JUN-1999	9905-0140901
PR	30-JUN-1999	9905-0141287
PR	01-JUL-1999	9905-0141682
PR	01-JUL-1999	9905-0142134
PR	02-JUL-1999	9905-0142055
PR	06-JUL-1999	9905-0142380
PR	08-JUL-1999	9905-0142803
PR	09-JUL-1999	9905-0142920
PR	12-JUL-1999	9905-0142977
PR	13-JUL-1999	9905-0143564
PR	14-JUL-1999	9905-0144005
PR	15-JUL-1999	9905-0144005
PR	16-JUL-1999	9905-0144005
PR	19-JUL-1999	9905-0144326
PR	19-JUL-1999	9905-0144332
PR	19-JUL-1999	9905-0144332
PR	19-JUL-1999	9905-0144333
PR	19-JUL-1999	9905-0144334
PR	19-JUL-1999	9905-0144334
PR	20-JUL-1999	9905-0144335
PR	20-JUL-1999	9905-0144332
PR	20-JUL-1999	9905-0144362
PR	20-JUL-1999	9905-0144364
PR	21-JUL-1999	9905-0144384

[illegible]

PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 25-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 26-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 89.3%; Score 25; DB 21; Length 1238;
 Best Local Similarity 83.3%; Pred. No. 1.4e+03;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6
 |||||
 Db 317 eelink 322

RESULT 12

AAW95359
 ID AAW95359 standard; Protein; 265 AA.

XX AAW95359;

XX 26-MAR-1999 (first entry)

DE Canine Interleukin (IL)-1 alpha precursor polypeptide.

XX Interleukin; IL-1 alpha propeptide; apoptosis; variant; malignant;

KW cancer; cell death.

OS Canis sp.

XX W09854205-A1.

XX 03-DEC-1998.

XX 28-MAY-1998; 98WO-US10839.

XX 27-MAY-1998; 98US-0065647.

XX 30-MAY-1997; 97US-0048137.

XX (REGC) UNIV CALIFORNIA.

XX Lovett DH, Pollock AS, Turck J;

XX WPI; 1999-059804/05.

XX New polypeptides comprising an apoptosis-inducing domain and a

XX heterologous nuclear localisation sequence - useful for, e.g.

XX forming compositions and methods for selectively inducing apoptosis

XX in cancer cells

XX Disclosure; Page 65-66; 88pp; English.

XX The invention relates to a human interleukin (IL)-1 alpha propeptide

XX polypeptide. The polypeptide has actively in the selective induction of

XX apoptosis in cancer cells. The IL-1 alpha propeptide or its

XX variants which retain the apoptosis-inducing activity are used in

XX compositions and methods to selectively induce apoptosis in cancer cells,

XX especially malignant cancer cells. The polypeptides form the basis for

XX effective, apoptotic-based chemotherapy where apoptosis-inducing agents

CC selectively trigger cell death in cancerous cells, but do not harm

CC surrounding normal tissue. The present sequence represents a canine

CC IL-1 alpha precursor polypeptide. The IL-1 alpha propeptide polypeptides

CC of non-human origin can be derived from such precursor polypeptides.

XX Sequence 265 AA;

Query Match 85.7%; Score 24; DB 20; Length 265;
 Best Local Similarity 66.7%; Pred. No. 4.9e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6
 |||||
 Db 101 eelink 106

RESULT 13

AAW69825
 ID AAW69825 standard; Protein; 306 AA.

XX AAW69825;

XX 17-AUG-1995 (first entry)

DE Transposon Tn5401 resolvase.

XX Transposon; Tn5401; resolvase; insecticide; pesticide;

XX biological control agent; site-specific recombination; B.t.

XX Bacillus thuringiensis var. morrisoni.

XX W09502058-A.

XX 19-JAN-1995.

XX 08-JUL-1994; 94WO-US07886.

XX 08-JUL-1993; 93US-0089986.

XX 24-JUN-1994; 94US-0266408.

XX (ECOG-) ECOGEN INC.

XX Baum JA;

XX WPI; 1995-066902/09.

XX N-PSDB; AA085492.

XX New transposon Tn5401 from Bacillus thuringiensis and fragments

XX of it - and related plasmids and transformed B. thuringiensis

XX useful as insecticides

XX Disclosure; Page 49-58; 106pp; English.

XX Tn5401 is present in B.t. var. morrisoni strain EG2158

XX (NRRL-B-18213). EG2158 was transformed with donor plasmid pEG911

XX and recombinant plasmids were analyzed. 2 Contained a 5 kb insert

XX (given in AA085492) including 2 ORFs encoding a resolvase (AAW69825)

XX and a transposase (AAW5354). Tn5401 is used in site-specific

XX recombination to construct improved B.t. strains for biological

XX control apn.

XX Sequence 306 AA;

Query Match 85.7%; Score 24; DB 16; Length 306;
 Best Local Similarity 66.7%; Pred. No. 5.6e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6
 |||||
 Db 172 edvirm 177

RESULT 14

AAW60180 standard; Protein: 306 AA.

AAW60180;

03-SEP-1998 (first entry)

B. thuringiensis transposon Tn5401 resolvase protein.

Transposon: Tn5401; Bacillus thuringiensis; insecticide; toxin;

insect control; recombinase/resolvase; Spodoptera; lepidoptera;

Bacillus thuringiensis.

US5776449-A.

07-JUL-1998.

20-SEP-1996; 96US-0717312.

20-SEP-1996; 96US-0717312.

08-JUL-1993; 93US-0089986.

24-JUN-1994; 94US-0266408.

(ECOG-) ECOGEN INC.

Baum JA:

WPI: 1998-397942/34.

N-PSDB; AAV34699.

Insecticidal recombinant Bacillus thuringiensis strains - contain

transposable elements for expression of recombinant insecticidal

proteins

Example 1: Fig 1A-L; 57pp; English.

This represents the resolvase protein encoded by tnp1 of the Bacillus

thuringiensis transposon Tn5401. The transposable element Tn5401

encodes 'recombinase/resolvase' and transposase gene products, which

enables a desired gene product (the toxin) to be transferred between

bacterial strains. This is used in a method for site-specific

recombination system for construction of recombinant B. thuringiensis

strains. These strains contain insecticidal toxin protein genes and are

free of DNA not native to B. thuringiensis. Insecticidal compositions

containing these B. thuringiensis strains are used in insect control

methods. The B. thuringiensis species used in the compositions produce

proteinaceous crystalline inclusions during sporulation, which are highly

toxic to specific insects, e.g. Spodoptera and lepidopteran species. The

plasmids used enable genes not native to B. thuringiensis to be

transferred to strains. Previous genetic transfer of plasmids relied on

conjugation, which does not always enable the foreign gene product to be

expressed in B. thuringiensis, and plasmids that do transfer, sometimes

carry undesirable genes.

Sequence 306 AA:

Query Match 85.7%; Score 24; DB 19; Length 306;
 Best Local Similarity 66.7%; Pred. No. 5.6e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDITMR 6
 11:1111
 Db 172 edv1mr 177

RESULT 15

AAV81702 standard; Protein: 314 AA.

AAV81702;

02-JUN-2000 (first entry)

Streptococcus pneumoniae protein sequence ID29.

Streptococcus pneumoniae infection; immunogen; antigen; diagnosis; AIDS;

bacterial pneumonia; asplenia; heart disease; lung disease; alcoholism;

kidney disease; diabetes; immunosuppressive disorder; otitis media;

pneumococcal septicemia; sinusitis; meningitis; therapy.

Streptococcus pneumoniae.

W0200006738-AZ.

10-FEB-2000.

27-JUL-1999; 99WO-GB02452.

27-JUL-1998; 98GB-0016336.

19-MAR-1999; 99US-0125329.

(MICR-) MICROBIAL TECHNIQUES LTD.

Le Page RWF, Wells JM, Hanniffy SB, Hansbro PM;

WPI: 2000-195301/17.

N-PSDB; AA91798.

Streptococcal proteins and polynucleotides useful for diagnosis,

treatment and prophylaxis of bacterial infections

Claim 1: Page 37; 76pp; English.

This sequence represents a Streptococcus pneumoniae protein of the

invention. The proteins (or their homologues, derivatives and/or

fragments) are useful as immunogens or antigens. Immunogenic or antigenic

compositions comprising the proteins are useful as vaccines and also in

diagnostic assays. The sequences are useful for the detection or

diagnosis of S. pneumoniae infection, by contacting a sample to be tested

with them. Agents capable of antagonising, inhibiting or interfering with

the function or expression of the protein or polypeptide are useful in

medical compositions in the treatment or prophylaxis of S. pneumoniae

infection. As the sequences can be used to treat S. pneumoniae infection,

they can be used to treat bacterial pneumonia, which has high rates in

young children, the elderly, and in patients with predisposing conditions

such as asplenia, heart, lung and kidney disease, diabetes, alcoholism,

or with immunosuppressive disorders, especially AIDS. They can also be

used to treat pneumococcal septicemia, otitis media, sinusitis, and

meningitis.

Sequence 314 AA:

Query Match 85.7%; Score 24; DB 21; Length 314;
 Best Local Similarity 66.7%; Pred. No. 5.6e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDITMR 6
 11:1111
 Db 145 edv1mr 150

Search completed: March 18, 2002, 09:48:57
 Job time: 366 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:49:41 ; Search time 37.77 Seconds

(without alignments)
3.575 Million cell updates/sec

Title: US-09-544-665-5

Perfect score: 28

Sequence: 1 EELIMR 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents, AA:*

- 1: /cgn2_6/prodata/2/1aa/5a_COMB.pep.*
- 2: /cgn2_6/prodata/2/1aa/5b_COMB.pep.*
- 3: /cgn2_6/prodata/2/1aa/5a_COMB.pep.*
- 4: /cgn2_6/prodata/2/1aa/5b_COMB.pep.*
- 5: /cgn2_6/prodata/2/1aa/PCTUS_COMB.pep.*
- 6: /cgn2_6/prodata/2/1aa/backfiles1.pep.*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	27	96.4	2763	3	US-08-496-944-2
2	25	89.3	668	1	US-08-205-018-2
3	25	89.3	859	1	US-08-395-580-2
4	25	89.3	859	5	PCT-US95-02792-2
5	24	85.7	263	4	US-09-085-305-9
6	24	85.7	306	1	US-08-089-986-2
7	24	85.7	306	1	US-08-478-585-2
8	24	85.7	306	1	US-08-717-312-2
9	24	85.7	306	5	US-08-266-408-2
10	24	85.7	306	5	PCT-US94-07886-2
11	24	85.7	490	2	US-08-687-916-25
12	24	85.7	490	4	US-09-138-614-25
13	24	85.7	1698	4	US-09-315-793-12
14	23	82.1	33	1	US-08-781-020-6
15	23	82.1	33	4	US-09-038-935-6
16	23	82.1	137	2	US-08-771-602D-46
17	23	82.1	348	1	US-08-843-521-6
18	23	82.1	348	2	US-08-953-041-2
19	23	82.1	348	4	US-09-159-417-2
20	23	82.1	348	4	US-09-012-871-6
21	23	82.1	379	3	US-08-840-204-3
22	23	82.1	390	1	US-08-121-714-6
23	23	82.1	390	1	US-08-477-108A-6
24	23	82.1	390	2	US-08-477-112-6
25	23	82.1	390	5	PCT-US93-08322-6
26	23	82.1	402	1	US-08-315-461-7
27	23	82.1	402	3	US-08-948-997-4

28	23	82.1	402	3	US-08-840-204-2	Sequence 2, Appl1
29	23	82.1	402	4	US-09-348-817A-4	Sequence 4, Appl1
30	23	82.1	402	4	US-09-026-408-3	Sequence 3, Appl1
31	23	82.1	486	3	US-08-889-841B-8	Sequence 8, Appl1
32	23	82.1	491	3	US-08-889-841B-10	Sequence 10, Appl1
33	23	82.1	496	3	US-08-889-841B-12	Sequence 12, Appl1
34	23	82.1	496	3	US-08-889-841B-16	Sequence 16, Appl1
35	23	82.1	498	3	US-08-889-841B-2	Sequence 2, Appl1
36	23	82.1	498	3	US-08-889-841B-5	Sequence 5, Appl1
37	23	82.1	508	4	US-08-472-240A-16	Sequence 16, Appl1
38	23	82.1	545	2	US-08-467-822-30	Sequence 30, Appl1
39	23	82.1	545	4	US-08-432-697-30	Sequence 30, Appl1
40	23	82.1	545	4	US-08-466-248-30	Sequence 30, Appl1
41	23	82.1	546	3	US-08-470-260-6	Sequence 6, Appl1
42	23	82.1	546	3	US-08-471-491-6	Sequence 6, Appl1
43	23	82.1	546	4	US-08-466-662-6	Sequence 6, Appl1
44	23	82.1	600	3	US-08-860-519-12	Sequence 12, Appl1
45	23	82.1	608	4	US-09-257-490-13	Sequence 13, Appl1

ALIGNMENTS

RESULT 1
US-08-496-944-2
Sequence 2, Application US/08496944
Patent No. 6040496

GENERAL INFORMATION:

APPLICANT: Law, Marcus D
TITLE OF INVENTION: Use of Translationally altered RNA to
TITLE OF INVENTION: Confer Resistance to Maltose Dwarf Mosaic Virus and Other
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY USA
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/496,944
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1814
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2763 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-496-944-2

Query Match 96.4% Score 27; DB 3; Length 2763;
Best Local Similarity 83.3% Pred. No. 3.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6
DB 2064 EELIMR 2069

RESULT 2

US-08-205-018-2
Sequence 2, Application US/08205018
Patent No. 5554523
GENERAL INFORMATION:
APPLICANT: Reddy, Usharani R.
TITLE OF INVENTION: No. 5554523el Protein Kinase, Nucleic Acid
TITLE OF INVENTION: Sequences Encoding the Same and Methods Related
TITLE OF INVENTION: Thereeto
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and
ADDRESSEE: No. 5554523rls
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/205,018
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Gaumont, Rebecca R.
REGISTRATION NUMBER: 35,152
REFERENCE/DOCKET NUMBER: CH-0488
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 668 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-205-018-2

Query Match 89.3%; Score 25; DB 1; Length 668;
Best Local Similarity 66.7%; Pred. No. 2.3e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EELMR 6
Db 409 EELVWR 414

RESULT 3
US-08-395-580-2
Sequence 2, Application US/08395580
Patent No. 5676945
GENERAL INFORMATION:
APPLICANT: Usharani R. Reddy, David Pleasure and the Children's
APPLICANT: Hospital of Philadelphia
TITLE OF INVENTION: No. 5676945el Protein Kinase, Nucleic Acid
TITLE OF INVENTION: Sequences Encoding the Same and Methods Related Thereeto
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and No. 5676945rls
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: U.S.A.
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch disk, 720 Kb
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/395,580
FILING DATE: herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/205,018
FILING DATE: 01-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Rebecca L. Ralph (formerly Gaumont)
REGISTRATION NUMBER: 35,152
REFERENCE/DOCKET NUMBER: CH-0488
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 859 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-395-580-2

Query Match 89.3%; Score 25; DB 1; Length 859;
Best Local Similarity 66.7%; Pred. No. 3.1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EELMR 6
Db 409 EELVWR 414

RESULT 4
PCT-US95-02792-2
Sequence 2, Application PC/TUS9502792
GENERAL INFORMATION:
APPLICANT: Usharani R. Reddy, David Pleasure and the Children's
APPLICANT: Hospital of Philadelphia
TITLE OF INVENTION: Novel Protein Kinase, Nucleic Acid
TITLE OF INVENTION: Sequences Encoding the Same and Methods Related Thereeto
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and Norris
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: U.S.A.
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch disk, 720 Kb
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02792
FILING DATE: herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/205,018
FILING DATE: 01-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Rebecca L. Ralph (formerly Gaumont)
REGISTRATION NUMBER: 35,152
REFERENCE/DOCKET NUMBER: CH-0488
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 859 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
PCT-US95-02792-2

Query Match 89.3%; Score 25; DB 5; Length 859;
Best Local Similarity 66.7%; Pred. No. 3.1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EETIMR 6
Db 409 EELVMR 414

RESULT 5
US-09-085-305-9
Sequence 9, Application US/09085305
Patent No. 6191269
GENERAL INFORMATION:
APPLICANT: Pollock, Allan
APPLICANT: Lovett, David H.
APPLICANT: Turk, Johana
TITLE OF INVENTION: Selective Induction of Apoptosis in
TITLE OF INVENTION: Malignant Cancer Cells by Delivery of N-Terminal
TITLE OF INVENTION: Interleukin-1-Alpha Pro-Piece Polypeptide
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bozicevic & Reed, LLP
STREET: 285 Hamilton Ave, Suite 200
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/085,305
FILING DATE: 29-MAY-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Francis, Carol L.
REGISTRATION NUMBER: 36,513
REFERENCE/DOCKET NUMBER: 6510/102051
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-327-3400
TELEFAX: 650-327-3231
TELEX:
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 265 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6191269e
US-09-085-305-9

Query Match 85.7%; Score 24; DB 4; Length 265;
Best Local Similarity 66.7%; Pred. No. 1.7e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EETIMR 6
Db 101 EETIMR 106

RESULT 6
US-08-089-986-2

Sequence 2, Application US/08089986
Patent No. 5441884
GENERAL INFORMATION:

APPLICANT: Baum, James A.
TITLE OF INVENTION: Bacillus thuringiensis Transposon Tn5401
TITLE OF INVENTION: and its use in a Site-Specific Recombination System For
TITLE OF INVENTION: Recombinant Bacillus thuringiensis Strain Development
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.
STREET: 1601 Market Street, 36th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19103

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/089,986
FILING DATE: 1993/07/08
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Egoif, Christopher
REGISTRATION NUMBER: 27633
REFERENCE/DOCKET NUMBER: 7205-47
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-757-1590
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 306 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-089-986-2

Query Match 85.7%; Score 24; DB 1; Length 306;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EETIMR 6
Db 172 EDVIMR 177

RESULT 7
US-08-478-585-2
Sequence 2, Application US/08478585
Patent No. 5650308
GENERAL INFORMATION:
APPLICANT: Baum, James A.
TITLE OF INVENTION: Bacillus thuringiensis Transposon Tn5401
TITLE OF INVENTION: and its use in a Site-Specific Recombination System For
TITLE OF INVENTION: Recombinant Bacillus thuringiensis Strain Development
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.
STREET: 1601 Market Street, 36th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/478,585
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/089,986
FILING DATE: 08-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: BOOLF, Christopher
REGISTRATION NUMBER: 27633
REFERENCE/DOCKET NUMBER: 7205-47
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-757-1590
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 306 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-478-585-2

Query Match 85.7%; Score 24; DB 1; Length 306;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELMR 6
DB 172 EDVIMR 177

RESULT 8
US-08-717-312-2
Sequence 2, Application US/08717312
Patent No. 5776449
GENERAL INFORMATION:
APPLICANT: Baum, James A.
TITLE OF INVENTION: Bacillus thuringiensis Transposon Tn5401
TITLE OF INVENTION: and its use in a Site-Specific Recombination System for
TITLE OF INVENTION: Recombinant Bacillus thuringiensis Strain Development
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Paulitch Schwarze Jacobs & Nadel c/o A.S. Nadel
STREET: 1601 Market Street, 36th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/717,312
FILING DATE: 20-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/266,408
FILING DATE: 24-JUL-1994
APPLICATION NUMBER: US 08/089,986
FILING DATE: 08-JUL-1993
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 306 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-717-312-2

Query Match 85.7%; Score 24; DB 1; Length 306;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELMR 6
DB 172 EDVIMR 177

RESULT 9
US-08-266-408-2
Sequence 2, Application US/08266408
Patent No. 5843744
GENERAL INFORMATION:
APPLICANT: Baum, James A.
TITLE OF INVENTION: Bacillus thuringiensis Transposon Tn5401
TITLE OF INVENTION: and its use in a Site-Specific Recombination System for
TITLE OF INVENTION: Recombinant Bacillus thuringiensis Strain Development
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Paulitch Schwarze Jacobs & Nadel c/o A.S. Nadel
STREET: 1601 Market Street, 36th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/266,408
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/089,986
FILING DATE: 08-JUL-1993
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 306 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-266-408-2

Query Match 85.7%; Score 24; DB 2; Length 306;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELMR 6
DB 172 EDVIMR 177

RESULT 10
PCT-US94-07886-2
Sequence 2, Application PC/TUS9407886
GENERAL INFORMATION:
APPLICANT: Baum, James A.
TITLE OF INVENTION: Bacillus thuringiensis Transposon Tn5401
TITLE OF INVENTION: and its use in a Site-Specific Recombination System for
TITLE OF INVENTION: Recombinant Bacillus thuringiensis Strain Development
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Paulitch Schwarze Jacobs & Nadel c/o A.S. Nadel
STREET: 1601 Market Street, 36th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/07886
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/089,986
FILING DATE: 08-JUL-1993
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 306 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-07886-2

Query Match 85.7%; Score 24; DB 5; Length 306;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6
Db 172 EDVIMR 177

RESULT 11
US-08-687-916-25
Sequence 25, Application US/08687916
Patent No. 5908972
GENERAL INFORMATION:
APPLICANT: HOUTZ, Robert L.
TITLE OF INVENTION: ISOLATED SPINACH
RIBULOSE-1,5-BISPHOSPHATE CARBOXYLASE/OXYGENASE LARGE
TITLE OF INVENTION: SUBUNIT N-METHYLTRANSFERASE AND METHOD OF INACTIVATING
RIBULOSE-1,5-BISPHOSPHATE CARBOXYLASE/OXYGENASE LARGE
TITLE OF INVENTION: SUBUNIT N-METHYLTRANSFERASE ACTIVITY
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/687,916
FILING DATE: 29-JUL-1996
CLASSIFICATION: 800
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/391,000
FILING DATE: 21-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Dadio, Susan M.
REGISTRATION NUMBER: 40,373
REFERENCE/DOCKET NUMBER: 028750-138
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 490 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-687-916-25

Query Match 85.7%; Score 24; DB 2; Length 490;
Best Local Similarity 66.7%; Pred. No. 3.1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6
Db 196 EVELIR 201

RESULT 12
US-09-138-614-25
Sequence 25, Application US/09138614
Patent No. 6245541
GENERAL INFORMATION:
APPLICANT: HOUTZ, Robert L.
TITLE OF INVENTION: ISOLATED SPINACH
RIBULOSE-1,5-BISPHOSPHATE CARBOXYLASE/OXYGENASE LARGE
TITLE OF INVENTION: SUBUNIT N-METHYLTRANSFERASE AND METHOD OF INACTIVATING
RIBULOSE-1,5-BISPHOSPHATE CARBOXYLASE/OXYGENASE LARGE
TITLE OF INVENTION: SUBUNIT N-METHYLTRANSFERASE ACTIVITY
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/138,614
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/687,916
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dadio, Susan M.
REGISTRATION NUMBER: 40,373
REFERENCE/DOCKET NUMBER: 028750-138
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 490 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-138-614-25

Query Match 85.7%; Score 24; DB 4; Length 490;
Best Local Similarity 66.7%; Pred. No. 3.1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6
Db 196 EVELIR 201

RESULT 13
US-09-315-793-12
Sequence 12, Application US/09315793
Patent No. 6221597
GENERAL INFORMATION:
APPLICANT: Roberts, Christopher J.


```

; TITLE OF INVENTION: ESSENTIAL GENES OF YEAST AS TARGETS FOR ANTIFUNGAL
; TITLE OF INVENTION: AGENTS, HERBICIDES, INSECTICIDES AND ANTI-PROLIFERATION
; TITLE OF INVENTION: DRUGS
; FILE REFERENCE: 9301-048
; CURRENT APPLICATION NUMBER: US/09/315,793
; CURRENT FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 1698
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; US-09-315-793-12

Query Match      85.7%; Score 24; DB 4; Length 1698;
Best Local Similarity 83.3%; Pred. No. 1e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EELIM 6
    |||||
Db 494 EELIM 499

RESULT 14
; US-08-781-020-6
; Sequence 6, Application US/08781020
; Patent No. 5792749
; GENERAL INFORMATION:
; APPLICANT: Wright, H. Tonle
; TITLE OF INVENTION: METHOD AND COMPOSITION FOR LOWERING LOW
; TITLE OF INVENTION: DENSITY LIPOPROTEIN CHOLESTEROL
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WHITHAM, CURTIS, WHITHAM & MCGINN
; STREET: Reston Intl. Center, 11800 Sunrise Valley Dr.,
; STREET: Suite 900
; CITY: Reston
; STATE: VA
; COUNTRY: USA
; ZIP: 20191
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,020
; FILING DATE: 09-JAN-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Whitham, Michael E.
; REGISTRATION NUMBER: 32,635
; REFERENCE/DOCKET NUMBER: VCU-16
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-391-2510
; TELEFAX: 703-391-9035
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-781-020-6

Query Match      82.1%; Score 23; DB 1; Length 33;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EELIM 5
    |||||

```

```

Db 4 EELIM 8

RESULT 15
; US-09-038-935-6
; Sequence 6, Application US/09038935
; Patent No. 6150332
; GENERAL INFORMATION:
; APPLICANT: Janciuskiene, Sabina
; APPLICANT: Wright, H. Tonle
; TITLE OF INVENTION: METHOD AND COMPOSITION FOR LOWERING LOW DENSITY
; TITLE OF INVENTION: LIPOPROTEIN CHOLESTEROL
; FILE REFERENCE: 294022Ba
; CURRENT APPLICATION NUMBER: US/09/038,935
; CURRENT FILING DATE: 1998-03-12
; EARLIER APPLICATION NUMBER: 08/781,020
; EARLIER FILING DATE: 1997-01-19
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-038-935-6

```

```

Query Match      82.1%; Score 23; DB 4; Length 33;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EELIM 5
    |||||
Db 4 EELIM 8

```

Search completed: March 18, 2002, 09:49:42
Job time: 226 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:50:32 ; Search time 41.98 Seconds
(without alignments)
10.887 Million cell updates/sec

Title: US-09-544-665-5

Perfect score: 28

Sequence: 1 EELMR 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : PIR_68:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	96.4	1757	2 T14318	ubiquitin-protein
2	26	92.9	157	2 H69019	conserved hypothet
3	26	92.9	263	2 S76994	hypothetical prote
4	26	92.9	498	2 B82622	anthranilate synth
5	26	92.9	886	2 S48371	hypothetical prote
6	25	89.3	149	2 B75001	hypothetical prote
7	25	89.3	192	2 U50749	superoxide dismuta
8	25	89.3	246	2 G65183	probable UDP-N-ace
9	25	89.3	251	2 T11549	ATP binding protei
10	25	89.3	287	2 E75202	dipeptide abc tran
11	25	89.3	289	2 S39814	DNA polymerase hom
12	25	89.3	312	2 S39823	lipid A biosynthes
13	25	89.3	318	2 G82350	recombination prot
14	25	89.3	520	2 F70350	protein kinase (EC
15	25	89.3	668	2 UC2363	NADPH--ferrithemop
16	25	89.3	671	2 A56592	probable retrovira
17	25	89.3	841	2 DB4513	serine/threonine p
18	25	89.3	888	2 A55318	dual leucine zippe
19	25	89.3	888	2 UCS399	protein F4N2.4 (lm
20	25	89.3	891	2 A96715	hyic protein - Ana
21	25	89.3	1108	2 T31335	hypothetical prote
22	25	85.7	75	2 C75121	hypothetical prote
23	24	85.7	256	2 S74821	envelope glycosyl
24	24	85.7	294	2 S60525	probable polyprote
25	24	85.7	294	2 E83022	site-specific reco
26	24	85.7	306	2 T40587	hypothetical prote
27	24	85.7	340	2 B84771	chortimate synthas
28	24	85.7	365	2 C64602	probable dipeptide
29	24	85.7	387	2 C72617	

30	24	85.7	400	2 A39254	inositol-1,4-bisph
31	24	85.7	445	2 A56024	GDP dissociation i
32	24	85.7	491	2 T03759	ribulose-bisphosp
33	24	85.7	495	2 S31493	env polyprotein -
34	24	85.7	496	2 T44307	lipote acetyl-tra
35	24	85.7	514	2 T49413	probable cytochrom
36	24	85.7	655	2 F69121	conserved hypothet
37	24	85.7	663	2 T03217	LJM domain protein
38	24	85.7	764	2 H82737	malate oxidoreduct
39	24	85.7	846	1 VCLJND	env polyprotein pr
40	24	85.7	853	2 S54384	envelope polyprote
41	24	85.7	855	1 VCLJZR	env polyprotein pr
42	24	85.7	861	2 B84885	probable transcrip
43	24	85.7	920	2 T41050	conserved hypothet
44	24	85.7	1353	2 T27404	hypothetical prote
45	24	85.7	1376	2 G00043	osteonidogen - hum

ALIGNMENTS

RESULT 1
T14318
ubiquitin-protein ligase E3-alpha - mouse
N:Alternate names: N-recognin E3-alpha
C:Species: Mus musculus (house mouse)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T14318
R:Kwon, Y.T., Reiss, Y., Fried, V.A., Hersheo, A., Yoon, J.K., Gonda, D.K., Sangen,
Proc. Natl. Acad. Sci. U.S.A. 95, 7898-7903, 1998
A:Title: The mouse and human genes encoding the recognition component of the N-end r
A:Reference number: Z17977; M01D:98318583
A:Accession: T14318
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1757 <RMO>
A:Cross-references: EMBL:AF061555; NID:g3170886; PID:g3170887; PIDN:AAC40165.1
C:Genetics:
A:Gene: ubi1
A:Map position: 2

Query Match 96.4%; Score 27; DB 2; Length 1757;
Best Local Similarity 83.3%; Pred. No. 2.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 EELMR 6
Db 765 EELMR 770

RESULT 2
H69019
conserved hypothetical protein MTH1148 - Methanobacterium thermoautotrophicum (strain
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: H69019
R:Smith, D.R.; Doucette-Stamm, L.A.; Delonghery, C.; Lee, H.; Dubois, J.; Aldredge,
; Olu, D.; Spadafora, R.; Vicalire, R.; Wang, Y.; Mierzbowski, J.; Gibson, R.; Jivani,
K. S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H. f
A:Reference number: A69000; M01D:98037514
A:Accession: H69019
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-157 <MTH>
A:Cross-references: GB:AE000884; GB:AE000666; NID:g2622242; PIDN:AAB85637.1; PID:g26
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH1148

Query Match 92.9%; Score 26; DB 2; Length 157;
Best Local Similarity 83.3%; Pred. No. 38;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6
DB 135 EELIMR 140

RESULT 3

576994 hypothetical protein - *Synechocystis* sp. (strain PCC 6803)

C:Species: *Synechocystis* sp.

A:Variety: PCC 6803

C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999

C:Accession: S76994

R:Kanehisa, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O.K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, D.A. Res. 27, 109-136, 1996

A>Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*

A:Reference number: S74322; MUID:97061201

A:Accession: S76994

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-263 <KAN>

A:Cross-references: EMBL:DB4005; GB:AB001339; NID:91001779; PIDN:BA10686.1; PID:0101133

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 92.9%; Score 26; DB 2; Length 263;
Best Local Similarity 83.3%; Pred. No. 66;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6
DB 153 EELIMR 158

RESULT 4

B82622 arachnolate synthase component I XF1914 (imported) - *Xylella fastidiosa* (strain 9a5c)

C:Species: *Xylella fastidiosa*

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000

C:Accession: B82622

R:anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequen

Nature 406, 151-157, 2000

A>Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.

A:Reference number: A82515; MUID:20365717

A>Note: for a complete list of authors see reference number A59328 below

A:Accession: B82622

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-498 <SIM>

A:Cross-references: GB:AE004011; GB:AE003849; NID:9107007; PIDN:AA84720.1; GSPDB:GN001

A:Experimental source: strain 9a5c

R:Simpton, A.J.G.; Reimach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvares, R.; Brites, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carrero, D.M.; Carier, H

as-Neto, E.; Docena, C.; El-Dorri, H.; Facincenti, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Frohm

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Kriger, J.E.; Krumm, E.E.; Laig

Chado, M.A.; Madeira, A.M.F.; Matsushima, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.

, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak

A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, R.M.; Silva Jr., M.A.; da Silveir

M.; Tsuchiko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A:Reference number: A59328

A:Contents: annotation

C:Genetics:

A:Gene: XF1914

C:Superfamily: arachnolate synthase component I

Query Match 92.9%; Score 26; DB 2; Length 498;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6
DB 300 EELIMR 305

RESULT 5

S48371 hypothetical protein YIL073c - yeast (*Saccharomyces cerevisiae*)

C:Species: *Saccharomyces cerevisiae*

C>Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 04-Mar-2000

C:Accession: S48371

R:Brown, D.; Bowman, S.

submitted to the EMBL Data Library, September 1994

A:Reference number: S48364

A:Accession: S48371

A:Molecule type: DNA

A:Residues: 1-886 <BRO>

A:Cross-references: GB:247047; EMBL:237997; NID:9603997; PID:9763273; GSPDB:GN00009;

C:Genetics:

A:Gene: YIL073c

A:Map position: 9L

C:Superfamily: *Saccharomyces cerevisiae* hypothetical protein YIL073c

Query Match 92.9%; Score 26; DB 2; Length 886;
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6
DB 255 EELIMR 260

RESULT 6

B75001 hypothetical protein PAB1019 - *Pyrococcus abyssi* (strain Orsay)

C:Species: *Pyrococcus abyssi*

C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000

C:Accession: B75001

R:anonymous, GenomeScope

submitted to the EMBL Data Library, July 1999

A:Description: *Pyrococcus abyssi* genome sequence: Insights into archaeal chromosome ;

A:Reference number: A75001

A:Accession: B75001

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-149 <KAW>

A:Cross-references: GB:AJ248288; GB:AL096836; NID:95458960; PIDN:CAB50448.1; PID:954;

A:Experimental source: strain Orsay

C:Genetics:

A:Gene: PAB1019

C:Superfamily: *Pyrococcus horikoshii* hypothetical protein PH0500

Query Match 89.3%; Score 25; DB 2; Length 149;
Best Local Similarity 83.3%; Pred. No. 64;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6
DB 74 EELIMR 79

RESULT 7

JS0749 superoxide dismutase (EC 1.15.1.1) (Fe) - *Legionella pneumophila*

C:Species: *Legionella pneumophila*

C>Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 20-Jun-2000
C:Accession: JS0749

R:Amemura-Maekawa, J.; Watanabe, H.
submitted to JIPID, September 1992

A:Reference number: JS0749

A:Accession: JS0749

A:Molecule type: DNA

A:Residues: 1-192 <AME>

A:Cross-references: DDBJ:D12922; NID:g216768; PIDD:BA02306.1; PID:g216769

A:Experimental source: strain Philadelphia-1

C:Function:

A:Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen

C:Superfamily: superoxide dismutase (hm)

C:Keywords: iron; metalloprotein; oxidoreductase

F:27/74,157,161/Binding site: iron (His, His, Asp, His) #status predicted

Query Match

Best Local Similarity 89.3%; Score 25; DB 2; Length 192;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6

Db 54 EELIMR 59

RESULT 8

665183 Probable UDP-N-acetyl-D-mannosaminuronic acid transferase - Escherichia coli (strain K-12)

N:Altemate names: hypothetical protein o246

C:Species: Escherichia coli

C>Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 08-Oct-1999

C:Accession: G65183; S30688

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; CC

A:Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617

A:Accession: G65183

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-246 <BLAT>

A:Cross-references: GB:AE000455; GB:U00096; NID:g2367282; PIDD:AACT6801.1; PID:g2367289;

A:Experimental source: strain K-12, substrain MG1655

R:Daniels, D.L.; Plunkett III, G.; Burland, V.; Blattner, F.R.

Science 257, 771-778, 1992

A:Title: Analysis of the Escherichia coli genome: DNA sequence of the region from 84.5 t

A:Reference number: S30660; MUID:92358234

A:Accession: S30668

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-119, 'AV', 122-177, 'X', 179-246 <DNAM>

A:Cross-references: EMBL:M87049; NID:g83656; PIDD:AA67594.1; PID:g148197

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1992

C:Genetics:

A:Gene: rffM

OY 1 EELIMR 6

Db 173 EELIMR 178

RESULT 9

T11549 ATP binding protein - Streptococcus gordonii

C:Species: Streptococcus gordonii

C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 02-Sep-2000

C:Accession: T11549

R:Kolenbrander, P.E.; Andersen, R.N.; Ganeshkumar, N.

Infect. Immun. 62, 4469-4480, 1994

A:Title: Nucleotide sequence of the Streptococcus gordonii PK488 coaggregation adhes

A:Reference number: 217283; MUID:95012638

A:Accession: T11549

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-251 <KOL>

A:Cross-references: EMBL:L11577; NID:g310629; PID:g310631

A:Experimental source: strain PK488

C:Superfamily: short-chain ATP-binding cassette proteins; ATP-binding cassette homo

F:29-221/Domain: ATP-binding cassette homology <ABC>

Query Match

Best Local Similarity 89.3%; Score 25; DB 2; Length 251;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6

Db 181 EELIMR 186

RESULT 10

E75202 dipeptide abc transporter, dipeptide-binding protein PAB0093 - Pyrococcus abyssi (s1

C:Species: Pyrococcus abyssi

C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000

C:Accession: E75202

R:Anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A:Description: Pyrococcus abyssi genome sequence: Insights into archaeal chromosome

A:Reference number: A75001

A:Accession: E75202

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-287 <KAW>

A:Cross-references: GB:AJ248283; GB:AL096836; NID:g5457433; PIDD:CAB49068.1; PID:g54

A:Experimental source: strain Orsay

C:Genetics:

A:Gene: oppC-1; PAB0093

C:Superfamily: oligopeptide permease protein oppB

Query Match

Best Local Similarity 89.3%; Score 25; DB 2; Length 287;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6

Db 118 EELIMR 123

RESULT 11

S39814 DNA polymerase homolog pol-r- rye mitochondrion (strain Halo)

C:Species: mitochondrion Secale cereale (rye)

A:Variety: strain Halo

C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 05-Nov-1999

C:Accession: S39814

R:Dojman, G.; Tuzdzynski, P.

Curr. Genet. 25, 59-65, 1994

A:Title: A DNA-polymerase-related reading frame (pol-r) in the mtDNA of Secale cere

A:Reference number: S39814; MUID:94363738

A:Accession: S39814

A:Molecule type: DNA

A:Residues: 1-289 <DOH>

A:Cross-references: EMBL:X74132; NID:g439279; PIDD:CAA52229.1; PID:g439280

A:Experimental source: strain Halo; etiolated shoots

C:Genetics:

A:Gene: pol-r

A:Genome: mitochondrion

C:Keywords: mitochondrion

Query Match 89.3%; Score 25; DB 2; Length 289;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EELIMR 6
 |||||:
 Db 136 EELIMK 141

RESULT 12

S39823
 DNA polymerase homolog pol-r - rye mitochondrion (strain B-P2)
 C:Species: mitochondrion Secale cereale (rye)
 A:Variety: strain B-P2
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 05-Nov-1999
 C:Accession: S39823
 R:Dojman, G.; Tudyanski, P.
 Curr. Genet. 25, 59-65, 1994
 A:Title: A DNA-polymerase-related reading frame (pol-r) in the mtDNA of Secale cereale.
 A:Reference number: S39814; MUID:94363738
 A:Accession: S39823
 A>Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-312 <DOH>
 A:Cross-references: EMBL:X74133; NID:9439281; PIDN:CA452230.1; PID:9439282
 A:Experimental source: strain B-P2; elicited shoots
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1993
 C:Genetics:
 A:Gene: pol-r
 A:Genome: mitochondrion
 C:Keywords: mitochondrion

Query Match 89.3%; Score 25; DB 2; Length 312;
 Best Local Similarity 83.3%; Pred. No. 1.4e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EELIMR 6
 |||||:
 Db 136 EELIMK 141

RESULT 13

G82350
 11pid A biosynthesis lauroyl acyltransferase VC0213 [Imported] - Vibrio cholerae (strain
 C:Species: Vibrio cholerae
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
 C:Accession: G82350
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
 Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F.
 I., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.W.
 Nature 406, 477-483, 2000
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A:Reference number: A82035; MUID:20406833
 A:Accession: G82350
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-318 <HEI>
 A:Cross-references: GB:AE004111; GB:AE003852; NID:99654614; PIDN:AAF93389.1; GSPDB:GN001
 A:Experimental source: serogroup O1, strain N16961, biotype El Tor
 C:Genetics:
 A:Gene: VC0213
 A:Map position: 1

Query Match 89.3%; Score 25; DB 2; Length 318;
 Best Local Similarity 83.3%; Pred. No. 1.4e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EELIMR 6
 |||||:
 Db 136 EELIMK 141

Db 287 EELIMR 292

RESULT 14

F70350
 recombinant protein RecN - Aquifex aeolicus
 C:Species: Aquifex aeolicus
 C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 21-Jan-2000
 C:Accession: F70350
 V:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.
 Nature 392, 353-358, 1998
 A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
 A:Reference number: A70300; MUID:98196666
 A:Accession: F70350
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-520 <AOF>
 A:Cross-references: GB:AE00695; NID:92983180; PIDN:AAC06789.1; PID:92983189; GB:AE0
 A:Experimental source: strain VF5
 C:Genetics:
 A:Gene: recN
 C:Superfamily: recN protein

Query Match 89.3%; Score 25; DB 2; Length 520;
 Best Local Similarity 83.3%; Pred. No. 2.4e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EELIMR 6
 |||||:
 Db 66 EELIMR 71

RESULT 15

JC2363
 protein kinase (EC 2.7.1.37) ZPK - human
 N:Alternate names: leucine-zipper protein kinase
 C:Species: Homo sapiens (man)
 C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 02-Feb-2001
 C:Accession: JC2363
 R:Reddy, V.R.; Pleasure, D.
 Biochem. Biophys. Res. Commun. 202, 613-620, 1994
 A:Title: Cloning of a novel putative protein kinase having a leucine zipper domain f
 A:Reference number: JC2363; MUID:94311945
 A:Accession: JC2363
 A:Molecule type: mRNA
 A:Residues: 1-668 <RBD>
 A:Cross-references: EMBL:U07358
 A:Experimental source: brain
 A>Note: the nucleotide sequence for this amino acid sequence is inconsistent with th
 he codon ACC for residue 661 as pro, the codon GAGCACCTCCCA for residues 664-668
 C:Comment: This protein belongs to the family of non-receptor kinase.
 C:Genetics:
 A:Gene: GDB:ZPK
 A:Cross-references: GDB:383963; OMIM:600447
 A:Map position: 12q13-12q13
 C:Superfamily: unassigned ser/Thr or Tyr-specific protein kinases; protein kinase ho
 C:Keywords: ATP; leucine zipper; nucleotide binding; P-loop; phosphotransferase
 F:133-371/Domain: protein kinase homology <KIN>
 F:131-139/Region: protein kinase ATP-binding motif
 F:443-471/Region: leucine zipper motif
 F:538-545/Region: nucleotide-binding motif A (P-loop)
 F:152/Active site: Lys #status predicted

Query Match 89.3%; Score 25; DB 2; Length 668;
 Best Local Similarity 66.7%; Pred. No. 3.1e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EELIMR 6
 |||||:
 Db 409 EELIMR 414

Tue Mar 19 11:00:08 2002

us-09-544-665-5.rpr

Page 5

Search completed: March 18, 2002, 09:50:33
Job time: 242 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:52:27 ; Search time 24.53 Seconds
(without alignments)
8.968 Million cell updates/sec

Title: US-09-544-665-5
Perfect score: 28
Sequence: 1 EELMR 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	92.9	886	1	YIH3_YEAST
2	25	89.3	192	1	SODF_LEGPN
3	25	89.3	246	1	WECG_ECOLI
4	25	89.3	251	1	YSC1_STRGC
5	25	89.3	378	1	CLPH_ONCVO
6	25	89.3	520	1	RECNC_AOUAE
7	25	89.3	671	1	NCPR_MUSDO
8	25	89.3	1076	1	HSEB_CAVPO
9	24	85.7	154	1	ELYS_HAIKA
10	24	85.7	265	1	ILIA_CANFA
11	24	85.7	365	1	AROC_HELPY
12	24	85.7	400	1	INPP_BOVIN
13	24	85.7	445	1	GDIB_MOUSE
14	24	85.7	755	1	BDAL_RHIME
15	24	85.7	846	1	ENV_HVIND
16	24	85.7	853	1	ENV_HV122
17	24	85.7	855	1	ENV_HV126
18	24	85.7	863	1	ENV_HV128
19	24	85.7	1375	1	NRD2_HUMAN
20	23	82.1	65	1	CSRA_VIRCH
21	23	82.1	123	1	Y900_AOUAE
22	23	82.1	192	1	KADA_MERG
23	23	82.1	192	1	KADA_MERJA
24	23	82.1	192	1	KADA_METVO
25	23	82.1	197	1	GRPE_ECOLI
26	23	82.1	208	1	SEGD_BPT4
27	23	82.1	217	1	GRAP_HUMAN
28	23	82.1	264	1	OTF_DROME
29	23	82.1	269	1	MORI_PASMU
30	23	82.1	298	1	FIG1_YEAST
31	23	82.1	310	1	CB32_DROME
32	23	82.1	320	1	MDH_RHIV
33	23	82.1	320	1	MOAA_AOUAE

34	23	82.1	338	1	Y674_METJA	Q58087 methanococ
35	23	82.1	348	1	AROH_BUCAI	P57224 buchnera ap
36	23	82.1	348	1	AROH_BUCAP	P46245 buchnera ap
37	23	82.1	362	1	AROG_HAEIN	P44303 haemophilus
38	23	82.1	364	1	OTF_PARLI	076971 paracentrot
39	23	82.1	364	1	RM02_KIULA	P48535 kluyveromyce
40	23	82.1	367	1	CURD_STRCN	002587 streptomyce
41	23	82.1	368	1	FTZ1_ARCFU	029715 archaeoglob
42	23	82.1	371	1	RM02_YEAST	P12687 saccharomyce
43	23	82.1	400	1	PATL_MUSVI	P50448 muscicla vis
44	23	82.1	402	1	PATL_BOVIN	P13908 bos taurus
45	23	82.1	402	1	PATL_HUMAN	P05121 homo sapien

ALIGNMENTS

RESULT 1	YIH3_YEAST	STANDARD:	PRT:	886 AA.
ID	YIH3_YEAST			
AC	P40511:			
DT	01-FEB-1995 (Rel. 31, Created)			
DT	01-FEB-1995 (Rel. 31, Last sequence update)			
DE	20-AUG-2001 (Rel. 40, Last annotation update)			
DR	HYPOTHETICAL 102.4 KDA PROTEIN IN SER33-HOP1 INTERGENIC REGION.			
CN	YL073C.			
OS	Saccharomyces cerevisiae (Baker's yeast).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC	Saccharomycetales; Saccharomycetaceae; Saccharomyces.			
OX	NCBI_TaxID=4932;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=5286C / AB972;			
RA	Barrett B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,			
RA	Churcher C.M., Connor R., Copsey T., Dear S., Davell K., Jones A.,			
RA	Gentles S., Hamlyn N., Hornsall T.S., Hunt S., Javels K., James M.,			
RA	Louis E., Lye G., Moule T., Odell C., Pearson D.,			
RA	Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,			
RA	Walsh S.V., Whitehead S.;			
RL	Submitted (Dec-1994) to the EMBL/Genbank/DBJ databases.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb.ch/announce/			
CC	or send an email to license@isb.ch).			
CC	-----			
DR	EMBL: 237997; CA86097.1; -			
DR	PIR: S48371; S48371.			
DR	SGD: S0001335; SPO22.			
KW	HYPOTHETICAL protein.			
SO	SEQUENCE 886 AA; 102356 MW; 83084BC8DC38945 CRC64;			

Query Match 92.9%; Score 26; DB 1; Length 886;
Best Local Similarity 83.3%; Pred. No. 86;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 2
SODF_LEGPN STANDARD; PRT: 192 AA.
ID SODF_LEGPN
AC P31108;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE SUPEROXIDE DISMUTASE [FE] (EC 1.15.1.1).

```

GN SODB.
OS Legionella pneumophila.
OC Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group.
OC Legionellaceae; Legionella.
OX NCBI_TaxID=446;
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=97240958; PubMed=9086394;
RA Aiemura-Maekawa J., Kura F., Watanabe H.;
RT Cloning and nucleotide sequences of iron and copper-zinc superoxide
RT dismutase genes of Legionella pneumophila and their distribution
RT among Legionella species.
RL Jpn. J. Med. Sci. Biol. 49:167-186(1996).
CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS.
CC -1- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) = O(2) + H(2)O(2).
CC -1- SUBUNIT: HOMODIMER.
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D12922; BAA02306.1;
DR PIR: J50748; J50749.
DR HSP: P09157; J158.
DR InterPro: IP001189; SOD_MT.
DR Pfam: PF00081; sode1.1.
DR ProDom: PD000475; SOD_MT.1.
DR PROSITE: PS00088; SOD_MN.1.
KW Oxidoreductase; Iron.
FT METAL 27 27 IRON (BY SIMILARITY).
FT METAL 74 74 IRON (BY SIMILARITY).
FT METAL 157 157 IRON (BY SIMILARITY).
FT METAL 161 161 IRON (BY SIMILARITY).
SO SEQUENCE 192 AA; 21639 MW; CA30F5459C0B4E9E CRC64;

Query Match 89.3%; Score 25; DB 1; Length 192;
Best Local Similarity 83.3%; Pred. No. 30;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EELIMR 6
DB 54 EELIMR 59

RESULT 3
WECG_FCOLI
ID WECG_FCOLI STANDARD; PRT: 246 AA.
AC P27836;
DT 01-AUG-1992 (Rel. 23, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PROBABLE UDP-N-ACETYL-D-MANNOSAMINURONIC ACID TRANSFERASE (EC 2.4.1.-)
DE (UDP-MANNAKA TRANSFERASE).
WECG OR RFFM OR B3794.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
RX MEDLINE=9238234; PubMed=1379743;
RA Daniels D.L., Plunkett G. III, Burland V.D., Blattner F.R.;
RT Analysis of the Escherichia coli genome: DNA sequence of the region
RT from 84.5 to 86.5 minutes.

```

```

RL Science 257:771-778(1992).
RN
RP REVISIONS TO 120-121.
RX STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Colado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT The complete genome sequence of Escherichia coli K-12.
RL Science 277:1453-1474(1997).
CC -1- PATHWAY: INVOLVED IN THE SYNTHESIS OF ENTEROBACTERIAL COMMON
CC O-SIDE CHAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M87049; AA67594.1;
DR EMBL: AE00455; AAC76801.1;
DR PIR: S30688; S30688.
DR Ecogene: EG11458; wecg.
KW Lipopolysaccharide biosynthesis; Glycosyltransferase; Transferase;
KW Complete proteome.
FT CONFLICT 120 VL -> AV (IN REF. 1).
SO SEQUENCE 246 AA; 27928 MW; 0D9435DEA6A69CAC CRC64;

Query Match 89.3%; Score 25; DB 1; Length 246;
Best Local Similarity 83.3%; Pred. No. 40;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EELIMR 6
DB 173 QELIMR 178

RESULT 4
YSC1_STRGC
ID YSC1_STRGC STANDARD; PRT: 251 AA.
AC P42360;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN IN SCA 5' REGION
DE (ORF1).
OS Streptococcus gordonii challis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=29390;
RN
RP SEQUENCE FROM N.A.
RX STRAIN=PK488;
RX MEDLINE=95012638; PubMed=7927711;
RA Kolenbrander P.E., Andersen R.N., Ganeshkumar N.;
RT Nucleotide sequence of the Streptococcus gordonii PK488
RT coaggregation adhesin gene, scaa, and ATP-binding cassette.
RL Infect. Immun. 62:4469-4480(1994).
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

CC -----
 DR EMBL: L11577; AAA71945.1; .
 DR InterPro: IPR003593; AAA.
 DR InterPro: IPR003439; ABC_transport.
 DR InterPro: IPR001687; ATP_GTP_A.
 DR Pfam: PF00005; ABC_tran; 1.
 DR SMART: SM00382; AAA; 1.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
 DR Hypothetical protein; ATP-binding; transport.
 KW NP_BIND 46 53 ATP (POTENTIAL).
 FT REPEAT 251 AA; 28054 MW; AED59C66F677D031 CRC64;
 SO SEQUENCE

Query Match 89.3%; Score 25; DB 1; Length 251;
 Best Local Similarity 83.3%; Pred. No. 40;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EITLMR 6
 DB 181 EITLMR 186

RESULT 5
 CLPH.ONCVO STANDARD; PRT; 378 AA.
 ID CLPH.ONCVO
 AC P37801:
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE CALPONIN HOMOLOG O99M.
 OS Onchocerca volvulus.
 OC Onchocerca; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
 OC Onchocercidae; Onchocerca.
 OX NCBI_Taxid=6282;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE:95021519; PubMed-7935620;
 RA Irvine M., Hulme T., Prince A.M., Lustigman S.;
 RT Identification and characterization of an onchocerca volvulus cDNA
 RL clone encoding a highly immunogenic calponin-like protein.
 RL Mol. Biochem. Parasitol. 65:135-146(1994).
 CC -1- FUNCTION: COULD BE INVOLVED IN MUSCLE CONTRACTION.
 CC -1- TISSUE SPECIFICITY: FOUND IN THE LONGITUDINAL MUSCLES BELOW THE
 CC HYPODERMIS.
 CC -1- DEVELOPMENTAL STAGE: PRESENT IN INFECTIVE LARVAE AND ADULT STAGES.
 CC -1- MISCELLANEOUS: THERE ARE APPARENTLY SEVERAL FORMS OF THE NATIVE
 CC PROTEIN WHICH APPEAR TO BE EXPRESSED IN A STAGE-SPECIFIC MANNER.
 CC -1- SIMILARITY: BELONGS TO THE CALPONIN FAMILY.
 CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-22 IS THE INITIATOR.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U01099; AAA20215.1; ALT_SEQ.
 DR InterPro: IPR000557; Calponin_repeat.
 DR Pfam: PF00402; calponin; 7.
 DR PROSITE: PS01052; CALPONIN; 5.
 KW Antigen; Repeat.
 FT REPEAT 98 123 CALPONIN-LIKE 26 AA MOTIF.
 FT REPEAT 151 176 CALPONIN-LIKE 26 AA MOTIF.
 FT REPEAT 197 222 CALPONIN-LIKE 26 AA MOTIF.
 FT REPEAT 244 269 CALPONIN-LIKE 26 AA MOTIF.
 FT REPEAT 285 310 CALPONIN-LIKE 26 AA MOTIF.
 SO SEQUENCE 378 AA; 41852 MW; 565E891180BF591E CRC64;

Query Match 89.3%; Score 25; DB 1; Length 378;
 Best Local Similarity 66.7%; Pred. No. 63;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 OY 1 EITLMR 6
 DB 88 EITLMR 93

RESULT 6
 RECN.AQVAF STANDARD; PRT; 520 AA.
 ID RECN.AQVAF
 AC O66834;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE DNA REPAIR PROTEIN RECN (RECOMBINATION PROTEIN N).
 GN RECN OR AQ_561.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificales; Aquificaceae; Aquifex.
 OX NCBI_Taxid=63363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VF5;
 RX MEDLINE-98196666; PubMed-9537320;
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 RT aeolicus".
 RL Nature 392:353-358(1998).
 CC -1- FUNCTION: MAY BE INVOLVED IN RECOMBINATIONAL REPAIR OF DAMAGED
 CC DNA (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE RECN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE00695; AAC06789.1; .
 DR InterPro: IPR003439; ABC_transport.
 DR InterPro: IPR003405; SMC_C.
 DR Pfam: PF02483; SMC_C; 1.
 KW DNA repair; ATP-binding; Complete proteome.
 FT NP_BIND 29 36 ATP (POTENTIAL).
 SO SEQUENCE 520 AA; 60439 MW; A9708562ACBE901E CRC64;

Query Match 89.3%; Score 25; DB 1; Length 520;
 Best Local Similarity 83.3%; Pred. No. 88;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 EITLMR 6
 DB 66 EITLMR 71

RESULT 7
 NCPR.MUSDO STANDARD; PRT; 671 AA.
 ID NCPR.MUSDO
 AC O07994;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE NADPH-CYTOCHROME P450 REDUCTASE (EC 1.6.2.4) (CPR) (P450R).
 OS Musca domestica (House fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Muscoidae; Muscidae; Musca.
 OX NCBI_Taxid=7370;
 RN [1]

```

RP SEQUENCE FROM N.A.
RC STRAIN-RUTGERS; TISSUE=Abdomen;
RX MEDLINE-93284260; PubMed-8508186;
RA Koener J.F., Carino F.A., Feyerisen R.;
RT "The cDNA and deduced protein sequence of house fly NADPH-cytochrome
RT P450 reductase."
RL Insect Biochem. Mol. Biol. 23:439-447(1993).
CC -1- FUNCTION: THIS ENZYME IS REQUIRED FOR ELECTRON TRANSFER FROM NADP
CC TO CYTOCHROME P450 IN MICROSOMES. IT CAN ALSO PROVIDE ELECTRON
CC TRANSFER TO HEME OXIGENASE AND CYTOCHROME B5.
CC -1- CATALYTIC ACTIVITY: NADPH + 2 FERRICYTOCHROME = NADP(+) +
CC 2 FERROCYTOCHROME.
CC -1- COFACTOR: BINDS ONE MOLE EACH OF FAD AND FMN.
CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM. ANCHORED TO THE
CC ER MEMBRANE BY ITS N-TERMINAL HYDROPHOBIC REGION.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT DEVELOPMENT, WITH HIGH
CC LEVELS BEFORE METAMORPHOSIS AND LOW LEVELS IN PUPAE.
CC -1- SIMILARITY: THE FMN-BINDING DOMAIN IS RELATED TO BACTERIAL
CC FLAVODOXINS. THE FAD DOMAIN APPEARS HOMOLOGOUS WITH FERREDOXIN
CC NADP+ REDUCTASE, NADH CYTOCHROME B5 REDUCTASE, AND GLUTATHIONE
CC REDUCTASE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L19897; AAA29295.1; -.
DR HSSP: P00388; IAM0.
DR InterPro: IPR003097; FAD_binding.
DR InterPro: IPR001094; Flavoxm-like.
DR InterPro: IPR001709; Flavopyrid_cyt_reductase.
DR InterPro: IPR001433; Oxidored_fad.
DR Pfam: PF00667; FAD_binding_1.
DR Pfam: PF00175; Oxidored_fad; 1.
DR PRINTS: PR00369; FLAVODOXIN.
DR PRINTS: PR00371; FPNCR.
KW Oxidoreductase; Flavoprotein; FMN; FAD; NADP; Endoplasmic reticulum;
KM Membrane.
FT NP_BIND 167 198 FMN (PYRIMIDINE PART) (BY SIMILARITY).
FT NP_BIND 311 322 FAD (ADP PART) (BY SIMILARITY).
FT NP_BIND 448 458 FAD (FLAVIN PART) (BY SIMILARITY).
FT NP_BIND 524 542 NADP (RIBOSE PART) (BY SIMILARITY).
FT NP_BIND 618 634 NADP (ADP PART) (BY SIMILARITY).
SQ SEQUENCE 671 AA; 76358 MW; A00A2C35DF8D129 CRC64;

Query Match 89.3%; Score 25; DB 1; Length 671;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EELIMR 6
DB 235 EEWLMR 240

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Krubheffer M., Celin Y., Kaempf U., Forsmann W.-G.;
RL Submitted (AUG-1996) to the EMBL/Genbank/DDJ databases.
CC -1- FUNCTION: RECEPTOR FOR THE E. COLI HEAT-STABLE ENTEROTOXIN (E. COLI
CC ENTEROTOXIN MARKELY STIMULATES THE ACCUMULATION OF CGMP IN
CC MAMMALIAN CELLS EXPRESSING GC-C). ALSO ACTIVATED BY THE ENDOGENOUS
CC PEPTIDE GUANYLIN.
CC -1- CATALYTIC ACTIVITY: GTP -> 3',5'-CYCLIC GMP + PYROPHOSPHATE.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC FAMILY.
CC -1- SIMILARITY: CONTAINS 1 PROTEIN KINASE-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z74734; CA98989.1; -.
DR HSSP: Q02846; IAM1.
DR InterPro: IPR001828; ANF_receptor.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR001054; Guanylt_cyclase.
DR Pfam: PF01094; ANF_receptor; 1.
DR Pfam: PF00211; guanylate_cyc; 1.
DR Pfam: PF00069; pkinase; 2.
DR SMART: SM00044; CYCC; 1.
DR PROSITE: PS00457; GUANYLATE CYCLASES_1; 1.
DR PROSITE: PS00125; GUANYLATE CYCLASES_2; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW Receptor; Transmembrane; Glycoprotein; Lyase; CGMP synthesis; Signal.
FT SIGNAL 1 23
FT CHAIN 24 1076
FT DOMAIN 24 433
FT DOMAIN 434 457
FT DOMAIN 458 1076
FT DOMAIN 492 752
FT DOMAIN 827 957
FT CARBOHYD 35 35
FT CARBOHYD 82 82
FT CARBOHYD 191 191
FT CARBOHYD 198 198
FT CARBOHYD 287 287
FT CARBOHYD 306 306
FT CARBOHYD 310 310
FT CARBOHYD 348 348
FT CARBOHYD 360 360
FT CARBOHYD 405 405
SQ SEQUENCE 1076 AA; 123119 MW; 9B53F16D5E080E9B CRC64;

Query Match 89.3%; Score 25; DB 1; Length 1076;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EELIMR 6
DB 674 OEIMR 679

```

```

RESULT 8
ID HSER_CAVPO STANDARD; PRT; 1076 AA.
AC P70106;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HEAT-STABLE ENTEROTOXIN RECEPTOR PRECURSOR (GC-C) (INTESTINAL
DE GUANYLATE CYCLASE) (EC 4.6.1.2) (STA RECEPTOR) (GUANYLYL CYCLASE C).
GN GUCY2C OR GUC2C.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriocognathi; Caviidae; Cavia.
OC NCBI_Taxid=10141;

```

```

RESULT 9
ID ELYS_HALKA STANDARD; PRT; 154 AA.
AC 00182;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE EGG-LYSIN PRECURSOR (SPERM-LYSIN).

```

OS Hallois kantschakana (Pinto abalone).
 CC Eukaryota; Metazoa; Mollusca; Gastropoda; Archaeogastropoda;
 CC Halioidea; Halioidea;
 CC NCBI_Taxid=6457;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lee Y.H., Vacquier V.D.;
 RT "The divergence of species-specific abalone sperm lysins is promoted
 by positive Darwinian selection.";
 RL Biol. Bull. 182:97-104(1992).
 CC -1- FUNCTION: DISSOLVES THE EGG VITELLINE LAYER NONENZYMATICALLY
 DURING FERTILIZATION. IT CREATES A HOLE OF ABOUT 3 MU-M IN
 DIMETER THROUGH WHICH THE SPERM PASS.
 CC DIMER THROUGH WHICH THE SPERM PASS.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- TISSUE SPECIFICITY: SPERM.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: M59970; AAA29200.1; -
 CC HSSP: P04552; LLYN.
 DR InterPro: IPR001379; Egg_Lysin.
 DR Pfam: PF01303; Egg_Lysin; 1.
 DR ProDom: PD002750; Egg_Lysin; 1.
 KW Fertilization; Sperm; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 154 EGG-LYSIN.
 SQ SEQUENCE 154 AA; 1836 MW; 8875D2B9ECF3564 CRC64;
 Query Match 85.7%; Score 24; DB 1; Length 154;
 Best Local Similarity 83.3%; Pred. No. 43;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 EELIMR 6
 Db 39 EELIVR 44
 RESULT 10
 ID IL1A.CANFA STANDARD; PRT; 265 AA.
 AC 046612;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DE INTERLEUKIN-1 ALPHA PRECURSOR (IL-1 ALPHA).
 GN IL1A.
 OS Canis familiaris (Dog).
 CC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 CC NCBI_Taxid=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Straubinger A.F., Straubinger R.K.;
 RT Submitted (Feb-1998) to the EMBL/GenBank/DBJ databases.
 RL FUNCTION: PRODUCED BY ACTIVATED MACROPHAGES, IL-1 STIMULATES
 THYOCYTE PROLIFERATION BY INDUCING IL-2 RELEASE, B-CELL
 MATURATION & PROLIFERATION, & FIBROBLAST GROWTH FACTOR ACTIVITY.
 IL-1 PROTEINS ARE INVOLVED IN THE INFLAMMATORY RESPONSE, BEING
 IDENTIFIED AS ENDOGENOUS PROTEINS, AND ARE REPORTED TO STIMULATE
 THE RELEASE OF PROSTAGLANDIN AND COLLAGENASE FROM SINOVIAL CELLS.
 CC (BY SIMILARITY).
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -1- DOMAIN: THE SIMILARITY AMONG THE IL-1 PRECURSORS SUGGESTS THAT THE
 AMINO ENDS OF THESE PROTEINS SERVE SOME AS YET UNDEFINED FUNCTION.
 CC MISCELLANEOUS: THE LACK OF A SPECIFIC HYDROPHOBIC SEGMENT IN THE
 PRECURSOR SEQUENCE SUGGESTS THAT IL-1 IS RELEASED BY DAMAGED CELLS

OR IS SECRETED BY A MECHANISM DIFFERING FROM THAT USED FOR OTHER
 SECRETORY PROTEINS.
 CC -1- SIMILARITY: BELONGS TO THE IL-1 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AF047011; AAC03066.1; -
 CC InterPro: IPR000975; Interleukin_1.
 DR InterPro: IPR003502; Interleukin_1-prop.
 DR Pfam: PF00340; IL1; 1.
 DR Pfam: PF02394; IL1-prop; 1.
 DR PRINTS: PR00264; INTERLEUKIN1.
 DR SMART: SM00125; IL1; 1.
 DR PROSITE: PS00253; INTERLEUKIN_1; 1.
 KW Cytokine; Macrophage; Mitogen; Inflammatory response; Pyrogen.
 FT PROPP 1 108 BY SIMILARITY.
 FT CHAIN 109 265 INTERLEUKIN-1 ALPHA.
 FT CARBOHYD 98 98 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 137 137 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 265 AA; 30521 MW; CCID316CA3717386 CRC64;
 Query Match 85.7%; Score 24; DB 1; Length 265;
 Best Local Similarity 66.7%; Pred. No. 77;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 EELIMR 6
 Db 101 EELIVR 106
 RESULT 11
 ID AROC_HELPY STANDARD; PRT; 365 AA.
 AC P56122;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE CHORISMATE SYNTHASE (EC 4.6.1.4) (5-ENOLPHOSPHORYLSHIMIKIMATE-3-PHOSPHATE
 PHOSPHOLYASE).
 GN AROC OR HP0663.
 OS Helicobacter pylori (Campylobacter pylori).
 CC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 CC Helicobacter.
 CC NCBI_Taxid=210;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPRAIN-26695 / ATCC 700392;
 MEDLINE-97394467; PubMed-9252185;
 RA Tomb J.F., White O., Kerejavage A.R., Clayton R.A., Sutton G.G.,
 RA Fleischmann R.D., Ketchum J.J., Zhou L., Klenk H.-P., Gill S., Dougherty B.A.,
 RA Nelson K., Quackenbush J., Zhou L., Kirnke E.F., Peterson S.,
 RA Loftus B., Richardson D., Dodson R., Khakh H.G., Hickley E.K.,
 RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
 RA Cotton M.D., Melman J.M., Fujii C., Bowman C., Watney L., Wallin E.,
 RA Venter J.C.;
 RT "The complete genome sequence of the gastric pathogen Helicobacter
 pylori.";
 RL Nature 388:539-547(1997).
 CC -1- CATALYTIC ACTIVITY: 5-O-(1-CARBOXYVINYL)-3-PHOSPHOSHIMIKIMATE -
 CHORISMATE + ORPHOSPHATE
 CC -1- COFACTOR: REDUCED FLAVIN (BY SIMILARITY).
 CC -1- PATHWAY: SEVENTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
 THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
 CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE CHORISMATE SYNTHASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: A000579; AAD0726.1; -
 CC TIGR: HP0663; -
 CC InterPro: IPR000453; Chorismate_synth.
 CC Pfam: PF01264; Chorismate_synth. 1.
 CC Prodom: P002941; Chorismate_synth. 1.
 CC PROSITE: PS00787; CHORISMATE_SYNTHASE_1; 1.
 CC PROSITE: PS00788; CHORISMATE_SYNTHASE_2; 1.
 CC PROSITE: PS00789; CHORISMATE_SYNTHASE_3; 1.
 CC Lyase; Aromatic amino acid biosynthesis; Complete proteome.
 CC
 CC SEQUENCE 365 AA; 40100 MW; 98ABC708E1909561 CRC64;

Query Match 85.7%; Score 24; DB 1; Length 365;
 Best Local Similarity 83.3%; Pred. No. 1.1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6
 DB 287 EDITOR 292

RESULT 12
 INPP_BOVIN STANDARD; PRT; 400 AA.
 AC P21327;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE INOSITOL POLYPHOSPHATE 1-PHOSPHATASE (EC 3.1.3.57) (IPPASE) (IPP).
 GN INPP1.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCB1_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Brain.
 RX MEDLINE=91088555; PubMed=2175905;
 RA York J.D., Majerus P.W.;
 RT Isolation and heterologous expression of a cDNA encoding bovine
 RT inositol polyphosphate 1-phosphatase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:9548-9552(1990).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
 RA MEDLINE=95034747; PubMed=7947723;
 RA York J.D., Ponder J.W., Chen Z.W., Mathews F.S., Majerus P.W.;
 RT "Crystal structure of inositol polyphosphate 1-phosphatase at 2.3-A
 RT resolution.";
 RL Biochemistry 33:13164-13171(1994).
 CC -1- CATALYTIC ACTIVITY: D-MTO-INOSITOL 1,4-BISPHOSPHATE + H(2)O =
 CC D-MTO-INOSITOL 4-PHOSPHATE + ORTHOPHOSPHATE.
 CC -1- ENZYME REGULATION: INHIBITED BY Li(+).
 CC -1- PATHWAY: PHOSPHATIDYL INOSITOL SIGNALING PATHWAY.
 CC -1- SUBUNIT: MONOMER.
 CC -1- MISCELLANEOUS: ACTS ON INOSITOL 1,4-BISPHOSPHATE AND INOSITOL
 CC 1,3,4-TRIPHOSPHATE (FORMING INOSITOL 3,4-BISPHOSPHATE) WITH A 4
 CC TIMES HIGHER AFFINITY FOR THE BISPHOSPHATE.
 CC -1- SIMILARITY: BELONGS TO THE INOSITOL MONOPHOSPHATASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M55916; AAA30588.1; -
 CC PIR: A39254; A39254.
 CC DR PDB: 1IMP; 26-JAN-95.
 CC InterPro: IPR000760; Inositol_P.
 CC Pfam: PF00439; Inositol_P. 1.
 CC PRINTS: PR00378; INOSPHATASE.
 CC PROSITE: PS00629; IMP_1; 1.
 CC PROSITE: PS00630; IMP_2; 1.
 CC Hydrolyase; Lithium; 3D-structure.
 CC
 CC FT CONFID 84
 CC SEQUENCE 400 AA; 43931 MW; CF385265E4C674AA CRC64;

Query Match 85.7%; Score 24; DB 1; Length 400;
 Best Local Similarity 83.3%; Pred. No. 1.2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6
 DB 90 EDITOR 95

RESULT 13
 GDI1_MOUSE STANDARD; PRT; 445 AA.
 ID GDI1_MOUSE
 AC P50397;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE RAB GDP DISSOCIATION INHIBITOR BETA (RAB GDI BETA) (GDI-2).
 GN GDI2 OR RABGDI2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCB1_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/C; TISSUE=Skeletal muscle;
 RX MEDLINE=94217740; PubMed=7513052;
 RA Shisheva A., Suedhof T.C., Czech M.P.;
 RT "Cloning, characterization, and expression of a novel GDP
 RT dissociation inhibitor isoform from skeletal muscle.";
 RL Mol. Cell. Biol. 14:3459-3468(1994).
 CC -1- FUNCTION: REGULATES THE GDP/GTP EXCHANGE REACTION OF MOST RAB
 CC PROTEINS BY INHIBITING THE DISSOCIATION OF GDP FROM THEM, AND THE
 CC SUBSEQUENT BINDING OF GTP TO THEM.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC AND MEMBRANE-ASSOCIATED.
 CC -1- TISSUE SPECIFICITY: UBQUITOUSLY EXPRESSED.
 CC -1- SIMILARITY: BELONGS TO THE TCD/MRS6 FAMILY OF GDP DISSOCIATION
 CC INHIBITOR.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U07951; AAB16908.1; -
 CC HSRF: P21856; 1GND.
 CC MGD: MG1:99645; GDI2.
 CC InterPro: IPR002005; Rab_GDI_REP.
 CC Pfam: PF00996; GDI; 1.
 CC PRINTS: PR00891; RABGDI2.
 CC PRINTS: PR00892; RABGDI.
 CC GTPase activation.
 CC
 CC SEQUENCE 445 AA; 50512 MW; 7FFD92EA950BE49 CRC64;

Query Match 85.7%; Score 24; DB 1; Length 445;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6
 DB 256 EELIMQ 261

RESULT 14

ID_BCAL_RHIME STANDARD; PRT: 755 AA.

AC 059750;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE BETA-GALACTOSIDASE (EC 3.2.1.23) (LACTASE).
 GN LACT2.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 CC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 CC Rhizobiaceae; Sinorhizobium.
 CC NCBI_Taxid=382;
 RX [1]
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.

RC STRAIN-102F34;
 RX MEDLINE-94215895; PubMed-8163182;
 RA Fanning S., Leahy M., Sheehan D.;
 RT "Nucleotide and deduced amino acid sequences of Rhizobium meliloti
 102F34 lacZ gene: comparison with prokaryotic beta-galactosidases and
 human beta-glucuronidase.";
 RL Gene 141:91-96(1994).
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL, NON-REDUCING BETA-D-
 GALACTOSE RESIDUES IN BETA-D-GALACTOSIDES.
 CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF GLYCOSYL HYDROLASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: L20757; AAA26296.1;
 DR InterPro: IPR001649; Glyco_Hydro.2.
 DR Pfam: PF00703; Glyco_Hydro.2; 1.
 DR PROSITE: PS00719; GLYCOSYL_HYDROL_F2_1; 1.
 DR PROSITE: PS00608; GLYCOSYL_HYDROL_F2_2; FALSE_NEG.
 KW Hydrolyase; Glycosidase.
 FT ACT_SITE 382 382 PROTON DONOR (BY SIMILARITY).
 FT ACT_SITE 463 463 NUCLEOPHILE (BY SIMILARITY).
 FT SEQUENCE 755 AA; 84148 MW; 09E339A92CCDD366 CRC64;

Query Match 85.7%; Score 24; DB 1; Length 755;
 Best Local Similarity 66.7%; Pred. No. 2.4e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6
 DB 551 EELIMK 556

RESULT 15

ENV_HVIND STANDARD; PRT: 846 AA.

AC P18799;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE ENVELOPE PROTEIN GP160 PRECURSOR [CONTAINS: EXTERIOR MEMBRANE

DE GLYCOPROTEIN (GP120); TRANSMEMBRANE GLYCOPROTEIN (GP41)).
 GN ENV.
 OS Human immunodeficiency virus type 1 (NDK isolate) (HIV-1).
 OC Viruses; Retrod viruses; Retroviridae; Lentivirus.
 OC NCBI_Taxid=11695;
 RX [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE-90034200; PubMed-2806917;
 RA Spire B., Sire J., Zacher V., Rey F., Barre-Sinoussi F., Galibert F.,
 RA Hampe A., Chermann J.C.;
 RT "Nucleotide sequence of HIV-1-NDK: a highly cytopathic strain of the
 RT human immunodeficiency virus.";
 RL Gene 81:275-284(1989).
 CC -1- MISCELLANEOUS: NDK, ISOLATED FROM A ZAIRIAN PATIENT AFFECTED WITH
 CC AIDS, AND IS A HIGHLY CYTOPATHOGENIC STRAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: M27323; AAA44873.1;
 DR PIR: J00066; VCLJND.
 DR HIV: M27323; ENVSDNDK.
 DR InterPro: IPR000328; ENV_GP41.
 DR InterPro: IPR000777; GP120.
 DR Pfam: PF00516; GP120.1.
 DR Pfam: PF00517; GP41.1.
 KW AIDS; Coat protein; Glycoprotein; Transmembrane;
 KW Signal.

FT SIGNAL 1 29
 FT CHAIN 30 501
 FT DISULFID 502 846
 FT DISULFID 53 73
 FT DISULFID 118 200
 FT DISULFID 125 191
 FT DISULFID 130 152
 FT DISULFID 213 242
 FT DISULFID 223 234
 FT DISULFID 291 328
 FT DISULFID 374 435
 FT DISULFID 381 408
 FT CARBOHYD 87 87
 FT CARBOHYD 129 129
 FT CARBOHYD 151 151
 FT CARBOHYD 179 179
 FT CARBOHYD 182 182
 FT CARBOHYD 229 229
 FT CARBOHYD 236 236
 FT CARBOHYD 257 257
 FT CARBOHYD 271 271
 FT CARBOHYD 284 284
 FT CARBOHYD 290 290
 FT CARBOHYD 351 351
 FT CARBOHYD 382 382
 FT CARBOHYD 388 388
 FT CARBOHYD 392 392
 FT CARBOHYD 395 395
 FT CARBOHYD 401 401
 FT CARBOHYD 438 438
 FT CARBOHYD 451 451
 FT CARBOHYD 452 452
 FT CARBOHYD 452 452
 FT CARBOHYD 601 601
 FT CARBOHYD 606 606
 FT CARBOHYD 615 615
 FT CARBOHYD 627 627

FT 846 AA; 96476 MW; 8A3B9DA527DE2B83 CRC64;
 SO SEQUENCE

Query Match

85.7%; Score 24; DB 1; Length 846;

Tue Mar 19 11:00:08 2002

us-09-544-665-5.rsp

Page 8

Best Local Similarity 83.3%; Pred. No. 2.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EELMR 6
||||:|
Db 263 EELMR 268

Search completed: March 18, 2002, 09:52:28
Job time: 297 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:51:59 ; Search time 78.82 Seconds
(without alignments)
11.135 Million cell updates/sec

Title: US-09-544-665-5

Sequence: 1 EELMR 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 14627329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :
1: SP archaea: *
2: SP bacteria: *
3: SP fungi: *
4: SP human: *
5: SP invertebrate: *
6: SP mammal: *
7: SP mhc: *
8: SP organelle: *
9: SP phage: *
10: SP plant: *
11: SP rodent: *
12: SP virus: *
13: SP vertebrate: *
14: SP unclassified: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	96.4	401	4 Q9H970	Q9H970 homo sapien
2	27	96.4	1757	11 Q70481	Q70481 mus musculu
3	26	92.9	157	1 Q27216	Q27216 methanobact
4	26	92.9	263	2 Q55923	Q55923 synchocyst
5	26	92.9	415	2 Q99278	Q99278 streptococ
6	26	92.9	498	2 Q9PC67	Q9PC67 xyella fas
7	25	89.3	71	12 Q99F59	Q99F59 human immun
8	25	89.3	81	12 Q73207	Q73207 human immun
9	25	89.3	99	12 Q9Q521	Q9Q521 human immun
10	25	89.3	99	12 Q9Q5V5	Q9Q5V5 human immun
11	25	89.3	102	12 Q9E4C3	Q9E4C3 human immun
12	25	89.3	107	12 Q9YX32	Q9YX32 human immun
13	25	89.3	107	12 Q9E4F9	Q9E4F9 human immun
14	25	89.3	107	12 Q9E4F8	Q9E4F8 human immun
15	25	89.3	107	12 Q9E4F5	Q9E4F5 human immun
16	25	89.3	107	12 Q9E4F4	Q9E4F4 human immun
17	25	89.3	114	12 Q9E4I0	Q9E4I0 human immun
18	25	89.3	125	12 Q9E4E6	Q9E4E6 human immun
19	25	89.3	149	1 Q9UYG4	Q9UYG4 pyrococcus

RESULT	ID	AC	DR	DE	OS	OC	OX	NCBI	SEQUENCE	Score	DB	Length	Best Local Similarity	Matches	Conservative	Mismatches	Indels	Gaps
20	25	89.3	150	12	Q9E4G5	Q9E4G5 human immun			Q9E4G5 human immun	96.44	401	1757	83.3%	5	1	0	0	
21	25	89.3	150	12	Q9E4C9	Q9E4C9 human immun			Q9E4C9 human immun	96.44	401	1757	83.3%	5	1	0	0	
22	25	89.3	158	12	Q9E4C8	Q9E4C8 human immun			Q9E4C8 human immun	96.44	401	1757	83.3%	5	1	0	0	
23	25	89.3	198	12	Q9E4G7	Q9E4G7 human immun			Q9E4G7 human immun	96.44	401	1757	83.3%	5	1	0	0	
24	25	89.3	198	12	Q9E4G2	Q9E4G2 human immun			Q9E4G2 human immun	96.44	401	1757	83.3%	5	1	0	0	
25	25	89.3	202	12	Q9E4H2	Q9E4H2 human immun			Q9E4H2 human immun	96.44	401	1757	83.3%	5	1	0	0	
26	25	89.3	202	12	Q9E4H1	Q9E4H1 human immun			Q9E4H1 human immun	96.44	401	1757	83.3%	5	1	0	0	
27	25	89.3	202	12	Q9E4E8	Q9E4E8 human immun			Q9E4E8 human immun	96.44	401	1757	83.3%	5	1	0	0	
28	25	89.3	202	12	Q9E4E7	Q9E4E7 human immun			Q9E4E7 human immun	96.44	401	1757	83.3%	5	1	0	0	
29	25	89.3	202	12	Q9E4E5	Q9E4E5 human immun			Q9E4E5 human immun	96.44	401	1757	83.3%	5	1	0	0	
30	25	89.3	202	12	Q9E4E4	Q9E4E4 human immun			Q9E4E4 human immun	96.44	401	1757	83.3%	5	1	0	0	
31	25	89.3	202	12	Q9E4E0	Q9E4E0 human immun			Q9E4E0 human immun	96.44	401	1757	83.3%	5	1	0	0	
32	25	89.3	202	12	Q9E4D9	Q9E4D9 human immun			Q9E4D9 human immun	96.44	401	1757	83.3%	5	1	0	0	
33	25	89.3	202	12	Q9E4D3	Q9E4D3 human immun			Q9E4D3 human immun	96.44	401	1757	83.3%	5	1	0	0	
34	25	89.3	202	12	Q9E4D4	Q9E4D4 human immun			Q9E4D4 human immun	96.44	401	1757	83.3%	5	1	0	0	
35	25	89.3	202	12	Q9E4D3	Q9E4D3 human immun			Q9E4D3 human immun	96.44	401	1757	83.3%	5	1	0	0	
36	25	89.3	202	12	Q9E4C4	Q9E4C4 human immun			Q9E4C4 human immun	96.44	401	1757	83.3%	5	1	0	0	
37	25	89.3	203	12	Q9E4G1	Q9E4G1 human immun			Q9E4G1 human immun	96.44	401	1757	83.3%	5	1	0	0	
38	25	89.3	203	12	Q9E4D8	Q9E4D8 human immun			Q9E4D8 human immun	96.44	401	1757	83.3%	5	1	0	0	
39	25	89.3	203	12	Q9E4C2	Q9E4C2 human immun			Q9E4C2 human immun	96.44	401	1757	83.3%	5	1	0	0	
40	25	89.3	204	12	Q9E4G4	Q9E4G4 human immun			Q9E4G4 human immun	96.44	401	1757	83.3%	5	1	0	0	
41	25	89.3	207	12	Q9E4H7	Q9E4H7 human immun			Q9E4H7 human immun	96.44	401	1757	83.3%	5	1	0	0	
42	25	89.3	207	12	Q9E4H4	Q9E4H4 human immun			Q9E4H4 human immun	96.44	401	1757	83.3%	5	1	0	0	
43	25	89.3	207	12	Q9E4F6	Q9E4F6 human immun			Q9E4F6 human immun	96.44	401	1757	83.3%	5	1	0	0	
44	25	89.3	207	12	Q9E4E9	Q9E4E9 human immun			Q9E4E9 human immun	96.44	401	1757	83.3%	5	1	0	0	
45	25	89.3	207	12	Q9E4D5	Q9E4D5 human immun			Q9E4D5 human immun	96.44	401	1757	83.3%	5	1	0	0	

ALIGNMENTS

RESULT 1
ID Q9H970
AC Q9H970
DR 01-MAR-2001 (TREMBLrel. 16, Created)
DE 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE CNA FL132968 FIS. CLONE NT2RP2005815.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isegal T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA Nakamura Y., Nagahari K., Masuno Y., Sasaki N.,
RT "NEDO human cDNA sequencing project";
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK023030; BAB14366.1; F9A68C7DADACBF3 CRC64;
SQ SEQUENCE 401 AA; 45391 MW; F9A68C7DADACBF3 CRC64;

Query Match 96.44; Score 27; DB 4; Length 401;
Best Local Similarity 83.3%; Pred. NO. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EELMR 6
||:||||
DB 306 EELMR 311

RESULT 2
ID Q70481
AC Q70481
DR 01-AUG-1998 (TREMBLrel. 07, Created)
DE 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE UBIQUITIN-PROTEIN LIGASE E3 COMMONEN N-RECOGNIN (UBIQUITIN-PROTEIN
LIGASE E3-ALPHA).

GN UBRI.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kwon Y.T., Reiss Y., Fried V.A., Hershko A., Yoon J.K., Gonda D.K.,
 RA Sangen P., Copeland N.G., Jenkins N.A., Varshavsky A.;
 RL Submitted (Apr-1998) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-446 FROM N.A.
 RC STRAIN=129/SVJ;
 RA Kwon Y.T., Reiss Y., Fried V.A., Hershko A., Yoon J.K., Gonda D.K.,
 RA Sangen P., Copeland N.G., Jenkins N.A., Varshavsky A.;
 RT *The mouse and human genes encoding the recognition component of the
 RT N-end rule pathway.*
 RL Submitted (May-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF061555; AAC23678.1; -
 DR EMBL: AF067379; AAC23678.1; -
 DR EMBL: AF067371; AAC23678.1; JOINED.
 DR EMBL: AF067372; AAC23678.1; JOINED.
 DR EMBL: AF067373; AAC23678.1; JOINED.
 DR EMBL: AF067374; AAC23678.1; JOINED.
 DR EMBL: AF067375; AAC23678.1; JOINED.
 DR EMBL: AF067376; AAC23678.1; JOINED.
 DR EMBL: AF067377; AAC23678.1; JOINED.
 DR EMBL: AF067378; AAC23678.1; JOINED.
 DR MGD: MGI:1277977; Ubr1.
 DR InterPro: IPR003126; 2nF_UBR1.
 DR Pfam: PF02207; 2nF_UBR1.1.
 DR SMART: SM00396; 2nF_UBR1.1.
 DR PROSITE: PS00732; RIBOSOMAL_S16; UNKNOWN.1.
 KW Ligase.
 SQ SEQUENCE 1757 AA; 200215 MW; 689FC06148258BEF CRC64;

Query Match 96.4%; Score 27; DB 11; Length 1757;
 Best Local Similarity 83.3%; Pred. No. 1e+03;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6
 Db 765 EELIMR 770

RESULT 3
 ID 027216 PRELIMINARY; PRT: 157 AA.
 AC 027216;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE CONSERVED PROTEIN.
 GN MTH1148.
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
 OC Methanothermobacter.
 OX NCBI_TaxID=145262;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DELTA H;
 RA MEDLINE=98037514; PubMed=9371463;
 RA Smith D.R., Doucette-Stamm L.A., Delonghery C., Lee H.-M., Dubois J.,
 RA Alredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lum K., Pochler B., Qiu D.,
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jirani N., Caruso A., Bush D., Safer H., Patel D., Prabhakar S.,
 RA McQuail S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
 RA Daniels C.J., Mao J.-I., Rice P., Nolling J., Reeve J.N.;
 RT *Complete genome sequence of Methanobacterium thermoautotrophicum
 RT delta: functional analysis and comparative genomics.*;

RL J. Bacteriol. 179:7135-7155(1997).
 DR EMBL: AE000884; AAB85637.1; -
 KW Complete proteome.
 SO SEQUENCE 157 AA; 17145 MW; 7EAF62F88BFF33AC CRC64;

Query Match 92.9%; Score 26; DB 1; Length 157;
 Best Local Similarity 83.3%; Pred. No. 1.4e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6
 Db 135 EELIMR 140

RESULT 4
 ID 055923 PRELIMINARY; PRT: 263 AA.
 AC 055923;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE HYPOTHETICAL 28.8 KDA PROTEIN.
 GN SLR0316.
 OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OX NCBI_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96127529; PubMed=8590279;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Sugura M., Tabata S.,
 RT *Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
 RT region from map positions 64% to 92% of the genome.*;
 RL DNA Res. 2:153-166(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97061201; PubMed=8905231;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hikosawa M., Sugita M., Sasamoto S., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
 RA Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
 RA Tabata S.;
 RT *Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions.*;
 RL DNA Res. 3:109-136(1996).
 DR EMBL: D64005; BAA10686.1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 263 AA; 28752 MW; 2FDE130FA3820763 CRC64;

Query Match 92.9%; Score 26; DB 2; Length 263;
 Best Local Similarity 83.3%; Pred. No. 2.4e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6
 Db 153 EELIMR 158

RESULT 5
 ID 099298 PRELIMINARY; PRT: 415 AA.
 AC 099298;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE HYPOTHETICAL PROTEIN SPY1012.
 GN SPY1012.
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;

OC Streptococcus.
 NX NCBI_TaxID-1314;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-SF370;
 RX MEDLINE-21192684; PubMed-11296296;
 RA Ferretti J.J., McShan M.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
 RA Primeaux C., Szate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
 RA Qian Y., Jia H.G., Nejar F.Z., Ren O., Zhu H., Song L., White J.,
 RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.,
 RT Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
 DR EMBL: AE006546; AAK3910.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 415 AA; 47745 MW; BABCT02BFC4FBE71 CRC64;

Query Match 92.9%; Score 26; DB 2; Length 415;
 Best Local Similarity 83.3%; Pred. No. 3.9e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EELMR 6
 DB 302 EELMR 307

RESULT 6
 ID 09P67 PRELIMINARY; PRT; 498 AA.
 AC 09P67;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE ANTHRANILATE SYNTHASE COMPONENT I.
 GN XPI1914.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 CC Xylella.
 NX NCBI_TaxID-2371;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-9A5C;
 RX MEDLINE-20365717; PubMed-10910347;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvaenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Biltons M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
 RA Pacicanti A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Praga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hohnselt J.D., Junqueira M.L., Kemper E.L., Kitzajima J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A., Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A., Jr., Pesquero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E., Jr., de Sa R.G., Sancelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva M.A. Jr.,
 RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
 RA Valada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.,
 RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
 RL Nature 406:151-159(2000).
 DR EMBL: AE004011; AAF84720.1;
 DR InterPro: IPR000350; Chorismate_bind.

DR Pfam: PF00425; Chorismate_bind. 2.
 DR PRINTS: PR00095; ANTSNHASEI.
 DR Prodom: PD00077; Chorismate_bind. 1.
 DR KW Complete proteome.
 SQ SEQUENCE 498 AA; 56109 MW; 945F54325856E927 CRC64;

Query Match 92.9%; Score 26; DB 2; Length 498;
 Best Local Similarity 83.3%; Pred. No. 4.7e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EELMR 6
 DB 300 EELMR 305

RESULT 7
 ID 099F59 PRELIMINARY; PRT; 71 AA.
 AC 099F59;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE ENVELOPE GLYCOPROTEIN (FRAGMENT).
 GN ENV.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 CC NCBI_TaxID-11676;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-BE2701;
 RA Candiotti D., Mundy C., Kadevele G., Nkhoma M., Bates I., Allain J.-P.,
 RT "Serologic and molecular screening for viruses in blood donors from
 RT Ntcheu, Malawi: high prevalence of HIV-1 subtype C and of markers of
 RT hepatitis B and C viruses.";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF309426; AAG49507.1;
 FT NON_TER 1 71
 FT NON_TER 1 71
 SQ SEQUENCE 71 AA; 7927 MW; 3DDBA8E683D55688 CRC64;

Query Match 89.3%; Score 25; DB 12; Length 71;
 Best Local Similarity 83.3%; Pred. No. 1.1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EELMR 6
 DB 9 EELMR 14

RESULT 8
 ID 073207 PRELIMINARY; PRT; 81 AA.
 AC 073207;
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE ENVELOPE GLYCOPROTEIN (FRAGMENT).
 GN ENV.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 CC NCBI_TaxID-11676;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-SAMPLE CANO01B;
 RA Roques P.,
 RT Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ006737; CA00211.1;
 DR InterPro: IPR000777; GP120.
 DR Pfam: PF00516; GP120. 1.
 KW Envelope protein. 1
 FT NON_TER 1 1

SO SEQUENCE 81 AA; 9048 MW; 0F4B3783F3AEAF99 CRC64;

Query Match 89.3%; Score 25; DB 12; Length 81;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDITOR 6
DB 4 EDITOR 9

RESULT 9
O90S21

ID O90S21 PRELIMINARY; PRT; 99 AA.

AC O90S21;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE GLYCOPROTEIN 120 (FRAGMENT).

ENV

OS Human immunodeficiency virus type 1.

OC Viruses; Retrovirdae; Retroviridae; Lentivirus.

OX NCBI_TaxID=11676;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=95TIN132;

RA Auenant W., Ayuthaya P.I., Duangchanda S., Mukai T., Kurata T., Ikuta K.;

RT

RT "Highly variable sequences at env V3 region of HIV-1 distributing

among Thai carriers during 1995 to 1997."

RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AB031918; BAA81868.1;

DR InterPro: IPR000777; GP120.

DR Pfam: PF00516; GP120.1.

FT NON_TER

FT NON_TER

SO SEQUENCE 99 AA; 11227 MW; 8C1275D49E96982 CRC64;

Query Match

Best Local Similarity 89.3%; Score 25; DB 12; Length 99;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDITOR 6
DB 6 EDITOR 11

RESULT 10
O90SV5

ID O90SV5 PRELIMINARY; PRT; 99 AA.

AC O90SV5;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE GLYCOPROTEIN 120 (FRAGMENT).

ENV

OS Human immunodeficiency virus type 1.

OC Viruses; Retrovirdae; Retroviridae; Lentivirus.

OX NCBI_TaxID=11676;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=97RVT592;

RA Auenant W., Ayuthaya P.I., Duangchanda S., Mukai T., Kurata T., Ikuta K.;

RT

RT "Highly variable sequences at env V3 region of HIV-1 distributing

among Thai carriers during 1995 to 1997."

RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AB031954; BAA81904.1;

DR InterPro: IPR000777; GP120.

DR Pfam: PF00516; GP120.1.

FT NON_TER

FT NON_TER

FT NON_TER 99 99
SO SEQUENCE 99 AA; 10795 MW; 8B05742A4BD0C66D CRC64;

Query Match 89.3%; Score 25; DB 12; Length 99;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDITOR 6
DB 6 EDITOR 11

RESULT 11
O9EAC3

ID O9EAC3 PRELIMINARY; PRT; 102 AA.

AC O9EAC3;

DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE ENVELOPE GLYCOPROTEIN (FRAGMENT).

ENV

OS Human immunodeficiency virus type 1.

OC Viruses; Retrovirdae; Retroviridae; Lentivirus.

OX NCBI_TaxID=11676;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-PATIENT 8;

RA Leroux C.;

RT "HIV-1 shedding" to the EMBL/GenBank/DBJ databases.

RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF256462; AAL6123.1;

DR InterPro: IPR000777; GP120.

DR Pfam: PF00516; GP120.1.

FT NON_TER

FT NON_TER

SO SEQUENCE 102 AA; 11510 MW; D6CF793B0415608C CRC64;

Query Match

Best Local Similarity 89.3%; Score 25; DB 12; Length 102;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDITOR 6
DB 2 EDITOR 7

RESULT 12
O9YN32

ID O9YN32 PRELIMINARY; PRT; 107 AA.

AC O9YN32;

DT 01-MAY-1999 (TREMBLrel. 10, Created)

DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE ENVELOPE GLYCOPROTEIN GP120 (FRAGMENT).

ENV

OS Human immunodeficiency virus type 1.

OC Viruses; Retrovirdae; Retroviridae; Lentivirus.

OX NCBI_TaxID=11676;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-SUBTYPE E;

RA Sato H., Shino T., Kodaka N., Taniguchi K., Tomita Y., Kato K., Miyakuni T., Takebe Y.;

RT "Evolution and biological characterization of human immunodeficiency

vertical virus subtype E gp120 V3 sequences following horizontal and

vertical transmission in a single family."

RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AB014805; BAA74625.1;

DR InterPro: IPR000777; GP120.

DR Pfam: PF00516; GP120.1.

FT NON_TER

FT NON_TER

FT NON_TER 107 107
SQ SEQUENCE 107 AA; 12305 MW; 369FAF9429FF81B3 CRC64;

Query Match
Best Local Similarity 89.3%; Score 25; DB 12; Length 107;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6
DB 14 EELILR 19

RESULT 13

ID 09EAF9 PRELIMINARY; PRT: 107 AA.
AC 09EAF9;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE ENVELOPE GLYCOPROTEIN (FRAGMENT).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT 8;
RA Leroux C.;
RT "HIV-1 shedding."
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF256426; AAG16087.1;
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
FT NON_TER 1
SQ SEQUENCE 107 AA; 12065 MW; 1600A374D8DE6F51 CRC64;

Query Match
Best Local Similarity 89.3%; Score 25; DB 12; Length 107;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6
DB 2 EELILR 7

RESULT 14

ID 09EAF8 PRELIMINARY; PRT: 107 AA.
AC 09EAF8;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE ENVELOPE GLYCOPROTEIN (FRAGMENT).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT 8;
RA Leroux C.;
RT "HIV-1 shedding."
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF256427; AAG16088.1;
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
FT NON_TER 1
SQ SEQUENCE 107 AA; 12032 MW; 2E274D8528E39E9E CRC64;

Query Match 89.3%; Score 25; DB 12; Length 107;

Best Local Similarity 89.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6
DB 2 EELILR 7

RESULT 15

ID 09EAF5 PRELIMINARY; PRT: 107 AA.
AC 09EAF5;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE ENVELOPE GLYCOPROTEIN (FRAGMENT).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT 8;
RA Leroux C.;
RT "HIV-1 shedding."
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF256430; AAG16091.1;
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
FT NON_TER 1
SQ SEQUENCE 107 AA; 12065 MW; 1600A374D8DE6F51 CRC64;

Query Match
Best Local Similarity 89.3%; Score 25; DB 12; Length 107;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EELIMR 6
DB 2 EELILR 7

Search completed: March 18, 2002, 09:51:59
Job time: 288 sec

Tue Mar 19 11:00:09 2002

us-09-544-665-5.rpt

Page 6